

Malaria's watershed

Malaria's moment has come, but success in control, let alone eradication, demands a renewed commitment to basic research.

A Global Malaria Action Plan, announced at the UN Millennium Development Goals Malaria Summit in New York on 25 September, has the ambitious goals of both reducing the malaria burden and eradicating the disease entirely.

Eradication any time soon might seem hopelessly optimistic, given the failure so far to make a serious dent in the number of malaria deaths. But much of the 274-page plan makes good sense. It calls for scaling up the use of existing tools, such as bednets, drugs and spraying, to near universal coverage, and then sustaining this effort for decades. True, this won't come cheaply. Funds for control have already grown from US\$250 million annually in 2004 to an estimated US\$1.1 billion this year; the plan calls for increasing that to \$5 billion annually until at least 2020. Likewise, total spending on malaria-related research has risen from \$265 million in 2003 to \$422 million in 2007; the plan would see that figure double to between \$750 million and \$900 million annually until 2018. Whether donors will rise to the challenge is a big question, given current economic woes. Still, it is heartening that at the summit, donors from governments, industry and philanthropic organizations pledged US\$3 billion.

Striking the right balance between basic and applied research is also critical. For example, the sequencing in 2002 of the genome of *Plasmodium falciparum*, the main parasite that causes malaria, has stimulated the hunt for new drug and vaccine candidates. This week's issue of *Nature* sees the addition of two more parasite sequences: *P. vivax*, which is less deadly than *P. falciparum*, and *P. knowlesi*, which mainly infects monkeys (pages 751, 757 and 799). These new sequences show how much more there is to learn: more than half of *P. falciparum*'s encoding genes still have no known function.

Basic research is also needed to stay ahead of drug resistance in the parasite and insecticide resistance in mosquitoes, and to get a better understanding of natural infection in humans. One surprise from *P. falciparum*'s genome is evidence that it evades the human immune system mainly by genomic and gene-expression diversity. *Plasmodium* seems to have different metabolic and physiological states, and can reprogram its gene expression. These findings could alter the way

researchers think about both drug and vaccine development.

The malaria drug and vaccine pipelines are healthier now than they have been for decades, but they are in urgent need of new candidates and approaches. So it was welcome news when the Bill & Melinda Gates Foundation, already the largest donor in malaria research, announced at the UN summit that it would spend US\$168 million to develop next-generation malaria vaccines. Moreover, this initiative will include early-stage laboratory research — most of the Gates Foundation's funding has so far focused on translational and clinical research.

Many scientists would like to see the foundation fund even more basic research, but this cannot be the foundation's responsibility alone. Its support of translational work has

"The malaria drug and vaccine pipelines are healthier now than they have been for decades."

rejuvenated the field over the past decade, and has helped get tools into the field. In the process, the malaria research community has become excessively and undesirably dependent on this one entity. Other research organizations would do well to step up to the plate and match the Gates Foundation's spending with their own basic-research funds. That would also go some way to addressing what scientists say is an unfortunate consequence of the emphasis on translational research: that scientists entering malaria research are less likely to choose basic science.

Any massive increase in research funding means that the malaria community must think about how to coordinate research across funding agencies. It is encouraging that the main research funders and scientists are to sit down as a group — called MalERA — over the next year to thrash out a research agenda for eradication. One lesson of the malaria and human genome projects is that consortia are a key route to delivery, focusing resources wisely, and avoiding duplication and excessive bureaucracy. It is also essential that international research recognizes the maturity of the malaria research community in the poorer countries where the disease is endemic — they should be on board as equals and not, as is too often the case, afterthoughts. ■

The Red List still matters

And the IUCN has more to offer than just data on the nearly extinct.

The International Union for Conservation of Nature, better known as the IUCN, is officially venerable. At 60, it is the oldest global conservation organization. Indeed, its best-known product — the Red List, a compendium of species threatened with

extinction — may seem a bit outdated in 2008. After all, the lesson of ecology is that species don't exist in isolation. They evolve and persist because of their relationships with all the other species around them. Conservationists these days usually talk about ecosystems as the units of interest, rather than species.

This trend towards broader thinking has not been lost on the IUCN. This week in Barcelona, Spain, the union is holding one of its four-yearly meetings. On the agenda is the release of the latest version of the Red List, which the union has been keeping since 1963, and which now covers nearly 45,000 species. But the union has also

authored an assessment of mammals that looks for larger patterns. For example, it turns out that marine mammals, and land mammals of south and southeast Asia, are in the worst shape — precisely the kind of knowledge that could help scarce conservation dollars go further (see page 717).

The IUCN is even getting into the business of predicting which species will one day become threatened as a result of climate change. Species with specific environmental needs or that have problems dispersing are likely to be most affected. The union is beginning similar work on other threats in collaboration with the Zoological Society of London.

This week, the union announced a new Global Mediterranean Action Network to help coordinate research on various Mediterranean-like biomes (including California and the Cape of South Africa). And IUCN president, Valli Moosa, was set to announce his vision for how the extraterritorial ocean should be managed so that the ecosystems of the high seas do not remain a conservation-free zone. Meanwhile, the IUCN's various programmes are working on issues ranging from making hotels more environmentally friendly to researching what an effective 'payments for environmental services' legal regime would look like.

Still, the heart of the organization is the Red List. Whatever its flaws, most people agree that the list is an irreplaceable indicator of global environmental health, using a metric that feels intuitive to

most of us: how many extinctions have we caused in the past four years? It is no surprise that the drafters of the Millennium Development Goals chose the Red List to be an indicator of progress towards reducing biodiversity loss.

Of course, the number of species explicitly studied and measured is dwarfed by the number of species yet to be discovered. To address this, the IUCN is beginning to use a sampling method to estimate the state of play in less-well-studied groups of organisms, such as invertebrates, taking its mammal survey as a calibration point.

The IUCN deserves credit for continuing to invest in the list, even though it clearly understands that conserving biodiversity requires much more than the list alone. Yes, ecosystems are more than just the sum of their parts — but there may be no more visceral way to convince people that conservation is worthwhile than to point to a species that has nearly died out due to human fumbling. The fishing cat (*Prionailurus viverrinus*) of south-east Asia has almost vanished. If reading that causes a pang, good. In these tumultuous economic times, when people are still trying to work out how to value 'ecosystem services', and when a full understanding of ecosystems is still decades away, the emotional force of extinction is nothing to sniff at. ■

"In these tumultuous economic times, the emotional force of extinction is nothing to sniff at."

Beta blockers?

Proprietary data formats may be legally defensible but open standards can be a better spur for innovation.

A historian of science and computing, and a scholar whose PhD thesis was on "professionalization of cooking among domestic servants in eighteenth-century France", might seem unlikely characters to find at the centre of a multimillion-dollar lawsuit. But that is exactly what has happened in the suit brought against George Mason University (GMU) in Fairfax, Virginia, by Thomson Reuters, the company probably best known for its ISI science indicators.

Dan Cohen, director of GMU's Center for History and New Media, and Sean Takats, a GMU history professor, are also directors of Zotero: open-source software developed by the history centre that lets researchers organize and share their digital information iTunes style, whether it is in the form of citations, documents or web pages. Zotero is free and popular, and has attracted some 1 million downloads since its launch in October 2006.

Thomson makes the proprietary bibliography software EndNote, and claims that Zotero is causing its commercial business "irreparable harm" and is wilfully and intentionally destroying Thomson's customer base. In particular, Thomson is demanding that GMU stop distributing the newer beta-version of Zotero that allegedly allows EndNote's proprietary data format for storing journal citation styles to be converted into an open-standard format readable by Zotero and other software. Thomson claims that Zotero "reverse engineered or decompiled" not only the format, but also the EndNote software itself.

The company is seeking a minimum of US\$10 million in damages annually until GMU halts distribution of Zotero's new feature. It also demands that GMU "terminate" the ability of each Zotero user to use or distribute any open-source files converted from EndNote's own data format. GMU seems ready to fight the suit; a spokesperson told *Nature* that the university believes it is "well within its rights", but declined to go into further detail given the ongoing litigation. Thomson was contacted but declined to comment, saying: "It is the policy of Thomson Reuters that we do not comment on pending litigation."

Thomson is claiming on the grounds that GMU has a site licence to EndNote, and that Zotero's actions breach the terms of the licensing contract. Thomson did not challenge GMU on grounds of copyright law, in which certain protections are in place to allow for creating interoperability. Thomson also claims that Zotero is infringing on the trademark 'EndNote' to induce Zotero users to convert EndNote's proprietary style files.

Litigation, which may go to a jury trial, is pending, so judging this case on its legal merits would be premature. But on a more general level, the virtues of interoperability and easy data-sharing among researchers are worth restating. Imagine if Microsoft Word or Excel files could be opened and saved only in these proprietary formats, for example. It would be impossible for OpenOffice and other such software to read and save these files using open standards — as they can legally do.

Competition between open-source and proprietary software is long-running, as personified by the struggle between Windows and Linux for desktop and server operating systems, but also in many branches of software used by scientists. Researchers tend to lean towards open sharing, but they will also pay for added-value features, and it's important that the playing field is level. Ultimately, the customer is king. ■

RESEARCH HIGHLIGHTS



Bird's eye view

Biol. Psychiatry doi:10.1016/j.biopsych.2008.06.012 (2008)

People with autism have incredibly keen eyesight, seeing almost as acutely as birds of prey.

The surprise finding from Emma Ashwin and her colleagues at the University of Cambridge, UK, shows that the unusually keen senses that have been associated with the condition since the 1940s stem not from how intensely autistic people feel their senses, but from how sharp their senses actually are.

Ashwin and her team tested 15 men with autism-spectrum disorders using the Freiburg Visual Acuity and Contrast Test, and found them to have, on average, 20:7 vision. This means they can see the same detail on an object 20 metres away that a person with average vision can see at 7 metres. Birds of prey have roughly 20:6 vision. What gives people with autism hawk-like vision isn't known.

F. JOREZ/GETTY

COSMOLOGY

Dark limit

J. Phys. A **41**, 412002 (2008)

Dark matter is a hypothetical class of particles that interact mainly through the force of gravity. How much dark stuff might be lurking around Earth is the subject of some debate, but Stephen Adler at the Institute for Advanced Study in Princeton, New Jersey, has set an upper limit.

Adler took data on the radius and period of a gravity-sensing satellite to estimate the planet's mass, including its dark matter, as well as lunar and asteroid orbital data to estimate the total mass inside the Moon's orbit.

Subtracting these two masses from a separate estimate of the combined masses of Earth and the Moon reveals the amount of dark matter within the Moon's orbit. He calculates that this can be no more than four billionths the mass of Earth, or 1.5×10^{15} kilograms.

BIOMECHANICS

Fungal ballistics

PLoS ONE **3**, e3237 (2008)

Some dung-feeding fungi squirt their spores at speeds of up to 25 metres per second. The spores need to travel several metres for herbivores to eat them — because most animals won't graze near their droppings — and thus to continue their life cycles.

The velocities were captured on high-speed video cameras (images right) by Nik Money at Miami University in Oxford, Ohio, and his colleagues. They showed that the pressure inside the squirt gun cell of several species is

similar to that of other fungal tissues. Their study will allow scientists to distinguish models that correctly describe the spore-ejection process.

ZOOLOGY

Boomerang bluefins

Science doi:10.1126/science.1161473 (2008)

Isotopic analysis of the ear bones of Atlantic bluefin tuna (*Thunnus thynnus*) has shown, for the first time, that an ocean-roaming fish returns to where it was born before spawning.

Jay Rooker of Texas A&M University in Galveston and his colleagues studied the ratio of oxygen-16 to oxygen-18 in the otoliths of these fish. Their findings indicate that 99.3% of the bluefin tuna spawning in the Gulf of Mexico and 95.8% of those spawning in the Mediterranean had swum back to their natal waters.

Populations of the fish have seen a precipitous decline from overfishing; the authors hope that their results will contribute to the species' future management.

PHYSICS

Light squeezing

Phys. Rev. Lett. **101**, 123601 (2008)

Physicists at the Pierre and Marie Curie University in Paris have proposed 'squeezing' light in order to measure the distance between objects in space more precisely.

The classic method, called the Einstein protocol, bounces pulses of light between two objects. But, at the quantum level, light is noisy, adding tiny measurement errors that can be significant when extreme accuracy is required.

Brahim Lamine and his colleagues calculate that squeezing light — shaping femtosecond laser pulses so as to reduce noise-inducing quantum fluctuations — might help. If their scheme works, it should provide greater control for positioning future flotillas of spacecraft such as Darwin, or space observatories like LISA — missions that aim to detect and observe Earth-like exoplanets and gravitational waves, respectively.

ORGANIC CHEMISTRY

State benefits

Angew. Chem. Int. Edn doi:10.1002/anie.200803648 (2008)

A new explosive that is among the most powerful known has the added benefits of being a solid at room temperature but having a low melting point of about 85 °C. This means it should make for a safer and more useful explosive than alternative esters of nitrate because it can be poured rather than pressed into shape.

The tetranitrate ester was made from a commercially available dioxane molecule by



© 2008 Macmillan Publishers Limited. All rights reserved

N. MONEY

David Chavez and his colleagues at the Los Alamos National Laboratory in New Mexico. They write that their invention reacted to impacts, sparks and friction like another high explosive, PTEN, which melts at more than 100 °C and is typically moulded as a solid. They predict that the new compound will detonate with as much force as the high-performance explosive HMX.

GEOSCIENCES

The melting ocean

Nature Geosci. doi:10.1038/ngeo316 (2008)
After 1997, a glacier that drains 7% of Greenland's ice sheet switched from thickening slowly to thinning quickly, causing the glacier's velocity to double. Several theories have been put forward to explain the change, including increased lubrication of the bedrock beneath the glacier. David Holland of New York University and his team conclude that it was induced by a sudden rise in the subsurface ocean temperature along Greenland's west coast.

They studied data from laser-altimeter surveys carried out by aircraft along 120 kilometres of the Jakobshavn glacier, and oceanographic observations recorded around the nearby port of Ilulissat. The pulse of warm water that arrived in 1997 came from the Irminger Sea near Iceland, they report, entering the subpolar gyre off Greenland after the North Atlantic Oscillation weakened during the winter of 1995–96.

PHYLOGEOGRAPHY

Viking mice

Proc. R. Soc. B doi:10.1098/rspb.2008.0958; 10.1098/rspb.2008.0959 (2008)

Mus musculus, the house mouse, has been colonizing new lands for several thousand years by hitchhiking with the humans whose crumbs it has come to rely on. Jeremy Searle of the University of York, UK, and his colleagues have used mouse mitochondrial DNA to retrace human migration.

They write that mice on the northern and western peripheries of the British Isles, particularly on the Orkney Islands, share a genetic lineage with Norwegian mice. These mice probably arrived with the Vikings — unlike mice from elsewhere in Britain, which are genetically more similar to German mice and probably reflect Iron Age migrations.

House mice on New Zealand, however, come from a mixture of countries, mirroring the complex history of migration to the archipelago from the late eighteenth century onwards. Before that, New Zealand was mouse-free.

MICROBIOLOGY

Half life

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0807707105 (2008)

The exceedingly abundant phytoplankton, *Emiliana huxleyi*, has unusual population dynamics. It can evade viral infection in its haploid form, when it has only one copy of each of its chromosomes, but it is susceptible to the same source of infection during the diploid part of its life cycle, when its cells contain twice as much DNA.

Miguel Frada at the Station Biologique in Roscoff, France, and his colleagues subjected the phytoplankton to giant phycodnaviruses. Unlike the diploid cells, the haploid ones did not burst open — perhaps owing to their uncalcified membranes somehow preventing the virus from entering the cells.



ZOOLOGY

Dik dik trick

Behav. Ecol. doi:10.1093/beheco/arn064 (2008)

Of the animals that understand other species' vocalizations, almost all are social creatures with complex calls of their own. But ecologists have identified an eavesdropper that is neither social nor particularly vocal: the dik-dik.

Daniel Blumstein and his colleagues at the University of California, Los Angeles, suspected that Gunther's dik-dik (*Madoqua guentheri*; pictured above), a heavily preyed miniature antelope, could benefit from eavesdropping. To find out whether it does, the researchers played alarm calls of the white-bellied go-away bird (*Corythaixoides leucogaster*) and non-alarmist calls from the slate-coloured boubou (*Laniarius funebris*) to a group of dik-diks at the Mpala Research Centre in Laikipia, Kenya.

The dik-diks in the study decreased their foraging and increased their head-turning only in response to the alarm calls.

W. BOLLMANN/PHOTOLIBRARY

JOURNAL CLUB

Ben Scheres
Utrecht University, The Netherlands

A plant scientist finds beauty in floral arrangements.

On the face of it, flower arranging is a fiddly affair, and its underlying rules are not immediately obvious to the beholder. But a plant's flowers are always arranged in one of three basic architectures, or 'inflorescences'. These take the form of panicles, loosely but highly branched clusters in which each flower has its own stalk (as in the foxglove); racemes, in which flowers are arranged individually along an unbranched, growing stem (the snapdragon); or cymes, typified by a cluster of branches at the end of a stem that each terminate with flower (the forget-me-not). Simple rules must lie behind this, and simple rules are the foodstuff of mathematical models.

That is the logic behind the work of Przemysław Prusinkiewicz at the University of Calgary in Alberta, Canada, and his colleagues. Last year, they published a model in which they imagined that meristems grow into shoots or flowers according to the value of a factor that they named 'veg' (P. Prusinkiewicz *et al. Science* **316**, 1452–1456; 2007). When veg is high, a shoot springs forth; when it is low, a blossom flourishes. Thus, if over time veg decreases at the same rate in all of a plant's growing tips, the model grows a panicle. Other simple rules give rise to a raceme or cyme.

Prusinkiewicz *et al.* found that, in *Arabidopsis*, a gene called *LEAFY* influences the value of veg. But how does this concept apply to plants with different architectures? Recently, Erik Souer of Vrije University in Amsterdam and his collaborators showed that modification of *LEAFY* activity is crucial for floral architecture in petunia, a cyme, just as the model predicts (E. Souer *et al. Plant Cell* **20**, 2033–2048; 2008). They identify a protein that activates *LEAFY* only in developing flower buds and that is essential for their architecture. I find the tidy simplicity of these findings more beautiful than any bouquet.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS

Nobel for AIDS virus discovery, finally

Two French virologists who discovered HIV share this year's Nobel Prize in Physiology or Medicine with a German oncovirologist who showed that human papilloma virus (HPV) causes cervical cancer. The Nobel Prize in Physics went to two Japanese researchers and a Japanese-born American for their work on symmetry breaking in subatomic physics.

Françoise Barré-Sinoussi and Luc Montagnier identified the human immunodeficiency virus (HIV) that causes AIDS in 1983, while working at the Pasteur Institute in Paris. They originally called it lymphadenopathy associated virus (LAV).

By awarding the prize to the French duo, the Nobel committee effectively ends years of bitter controversy arising from a counter claim for HIV's discovery by virologist Robert Gallo of the US National Institutes of Health in Bethesda, Maryland. In 1987 the French and US heads of state brokered an agreement to share the benefits of the discovery.

Many believed that the Nobel prize could not be awarded for this field of research, despite its



Françoise Barré-Sinoussi (left), Luc Montagnier (centre) and Harald zur Hausen share the medicine Nobel for their work on viruses.

importance, while tempers ran high. Now, the Nobel committee has made its position clear, saying the discovery of Barré-Sinoussi and Montagnier "was accepted by the research community and resulted in an explosion of scientific breakthroughs". It then refers to Gallo's "detection of a novel ... virus from a vast number of patients with AIDS or pre-AIDS in 1984 ... [which] showed considerable similarities with LAV-1". Gallo says he is "gratified" that he was considered "equally deserving", adding that he is proud of his colleagues.

The work of the French scientists has led to the development of diagnostic tools, blood

screening agents and drugs that have helped to fight spread of the disease and dramatically increased life expectancy.

Harald zur Hausen, former director of the German Cancer Research Center (DKFZ) in Heidelberg, Germany, was honoured for going against the prevailing dogma to discover that it was not a herpes simplex virus that causes cervical cancer, but in fact HPV.

"Virologists and gynaecologists thought his idea about the papilloma virus was very strange," says Herbert Pfister, a former colleague, now at the University of Cologne in Germany. "But he carried through his theory in a determinedly logical way, not caring about the controversy he was raising."

During the 1980s, zur Hausen described first HPV-16, which occurs in around half of human cervical cancers as well as other anogenital cancers, and then HPV-18, which accounts for a further 25% of cases. His work led to the development of vaccines against HPV, which are now in widespread use and expected to slash the rates of cervical cancer,

INSTITUT PASTEUR; U. ANDERSEN/GETTY IMAGES; REUTERS

Charitable bodies hit by credit crisis

The ongoing financial crisis in the United States and Europe is hitting major research charities and institutions.

Organizations such as the Wellcome Trust in London and Cold Spring Harbor Laboratory in New York say that their endowments have dropped for the first time in years. And some medical charities are bracing for a decline in big-dollar donations from corporations and wealthy individuals. "The events of the past few weeks are clearly going to have an impact on everybody," says Bruce Stillman, president of Cold Spring Harbor Laboratory.

Still, he and other leaders contacted by *Nature* are confident that sound financial management and careful planning will allow them to survive the economic crash.

The credit freeze, caused by bad loans in the housing sector in the United States and elsewhere, has brought down several leading banks and insurance companies. The time taken for US Congress to approve a \$700-

billion bailout package caused markets to fall further still, and governments have been bringing out emergency measures to shore up their economies.

The markets strongly influence trends in charitable giving, according to Patrick Rooney, director of research at Indiana University's Center on Philanthropy in Indianapolis. "When the stock market was growing in double digits in the late 1990s, philanthropy was growing in double digits as well," he says. But during the US recession of 2001–03, charitable giving dropped by a few percentage points each year.

Rooney says that it is probably too early to tell how much the crisis has affected charities and endowments. Nearly one-quarter of donations made by individuals and corporations are made during the autumn and early winter, and any decline in giving won't be felt until spring 2009, he says. He expects that organizations that depend mainly on endowments will fare better than

charities, so long as the economic downturn is not sustained. But if the economy slides into recession, he warns, "all bets are off".

Some charities have escaped the international economic downturn unscathed. The Howard Hughes Medical Institute's \$19-billion endowment is relatively unchanged compared with August 2007, according to Thomas Cech, the institute's outgoing president. Cech says that only about one-quarter of the endowment is invested in the market and that the losses in recent months have been offset by gains earlier in the year. "We are fully confident that we can meet our upcoming commitments," he says.

But others have been directly affected by the economic turmoil. The Wellcome Trust, the world's largest medical research charity, has seen its endowment drop by 5–10% to about £13 billion (US\$23 billion) this year, according to Mark Walport, the charity's director.

**HAVE YOUR SAY**

Comment on any of our news stories, online.

www.nature.com/news

which affects nearly 500,000 people a year and is the fifth most deadly cancer in women.

The physics prize went to Japan's Makoto Kobayashi of the High Energy Accelerator Research Organization (KEK) in Tsukuba and Toshihide Maskawa of Yukawa Institute for Theoretical Physics at Kyoto University for discovering the origin of the 'broken symmetry' that contributed to a preponderance of matter over antimatter in the Universe. They share the prize with Yoichiro Nambu of the Enrico Fermi Institute at the University of Chicago in Illinois, who discovered the mechanism by which spontaneous broken symmetry occurs in particle physics.

Symmetry breaking describes how symmetrical systems can suddenly show a preference for one direction over another. As a simple example, imagine balancing a pencil on its tip. Viewed from the top, the balanced pencil appears symmetrical, but after a time it will fall and point in a single direction.

This concept applies to many physical systems, including superconductors, but it was Nambu who extended the theory to fundamental particles, according to John Ellis, a theorist at CERN, Europe's particle-physics laboratory near Geneva, Switzerland. Physicists now think that symmetry breaking is responsible for the creation of the Higgs boson, the particle that is believed to endow all other particles with mass,

and which is a target of CERN's Large Hadron Collider.

Kobayashi and Maskawa showed how violation of symmetry could create more matter than antimatter in the Universe — a long-standing problem in particle physics. In the early 1970s, the pair showed that the interactions of quarks via a fundamental force, called 'the weak force', could cause 'CP-violation', a phenomenon by which some particles decay in a different way from their anti-matter counterparts.

"They wrote down this huge expression whose physical interpretation is the violation of symmetry between matter and antimatter," says Ken Peach, a physicist at the University of Oxford, UK. The equations also predicted a third family of quarks (the particles that make up protons and neutrons in an atom's nucleus).

The idea of a thirds family seemed "far-fetched" at the time, says Ellis, but Kobayashi and Maskawa's work has since been verified by two high-energy experiments. The Belle experiment at KEK and the BaBar experiment at the Stanford Linear Accelerator Center in California both measured the decay of particles made of bottom quarks. The physicists' predictions were borne out to a high degree of accuracy. ■

Alison Abbott and Geoff Brumfiel

To read coverage of the Nobel Prize in Chemistry, which had not been announced as *Nature* went to press, visit www.nature.com.

Still, Walport says that the Wellcome Trust is not anticipating any need to alter its spending. Like most charities, it gives only a small fraction of its endowment to research each year — around £500 million in 2007. That relatively low rate of spending combined with careful management ahead should allow the charity to continue normal operations. "We are weathering the storm as well as can be expected," Walport says.

Cold Spring Harbor's US\$300 million endowment is also down, says Stillman, by an estimated 5% as of 31 July. And the laboratory, which supports a broad range of basic biological science, depends on additional philanthropic and corporate support to fund its roughly \$120-million operating budget. Stillman says that the money often comes either directly from Wall Street firms or from wealthy investors. The latest fundraising efforts are on track, he says. But "obviously the situation will affect our income down the road".

Rooney says that corporations may be the first to cut back on their giving. "Corporate donations are largely driven by changes in profitability," he says.

Lean times could be especially difficult for charities that depend entirely on annual donations to fund research. "Donors may raise their bar," warns Deborah Brooks, co-founder of the Michael J. Fox Foundation for Parkinson's Research in New York. Brooks says she thinks that wealthy individual donors will probably continue to give large sums but "their first instinct is to give to fewer people".

Stillman says he expects that Wall Street's woes will have a "significant" effect on research universities and charities across the state. Indeed, one important source of funding has dried up for good. Lehman Brothers, which filed for bankruptcy last month, donated roughly \$7 million to biomedical research in 2007. Most of the money went to hospitals and research centres in the New York area, with the largest amount, \$6 million, going to Weill Cornell Medical College's Lehman Brothers Lung Cancer Research Center. A spokesperson for the school declined to comment on the centre's future, saying only that it was "grateful for the many years of support by Lehman Brothers". ■

Geoff Brumfiel

Kavli invests in institutes

US philanthropist Fred Kavli has embarked on a second round of grants of up to \$5 million to each of the 15 research institutes that were established in his name. The grants, Kavli told *Nature*, are conditional on the institutes finding gifts from other donors so that they can establish endowments of \$20 million.

Kavli, who made his \$600-million fortune through real estate and selling his company — which had become a leader in supplying aeronautic and automotive sensors — says that the second round could take five years or more. "Right now the market hasn't been very kind to us," he says. "We are hedging a little on this."

Seven years ago, when Kavli established the first of the institutes, at the University of California at Santa Barbara, he developed what became a relatively strict formula:



Fred Kavli is hoping to support 20 institutes.

on average each university got \$7.5 million, to be supplemented and used however — for a building, for operating expenses, or for endowed professorships.

Kavli got the name. John Carlstrom, director of the Kavli

Institute for Cosmological Physics at the University of Chicago in Illinois, says that his university was not the only one that thought the gift was relatively small to be exchanged for the naming rights to an entire institute. But many assumed, apparently correctly, that more money would be coming later.

The first of the second round of gifts was completed in May to the Kavli Institute for Particle Astrophysics & Cosmology to Stanford University in Menlo Park, California. A deal with the University of California at Santa Barbara is in the works as fundraisers find matching gifts.

Kavli, 81, says that he views the institutes, which specialize in nanoscience, neuroscience, astrophysics and theoretical physics, as his legacy — and as more important than the three \$1 million eponymous prizes that were awarded for the first time this year. But he says that he is not done yet; he eventually envisions 20 institutes. ■

Eric Hand

S. ANDERSEN/NORDISK FILM



HAVE YOUR SAY

Comment on any of our news stories, online.

www.nature.com/news

which affects nearly 500,000 people a year and is the fifth most deadly cancer in women.

The physics prize went to Japan's Makoto Kobayashi of the High Energy Accelerator Research Organization (KEK) in Tsukuba and Toshihide Maskawa of Yukawa Institute for Theoretical Physics at Kyoto University for discovering the origin of the 'broken symmetry' that contributed to a preponderance of matter over antimatter in the Universe. They share the prize with Yoichiro Nambu of the Enrico Fermi Institute at the University of Chicago in Illinois, who discovered the mechanism by which spontaneous broken symmetry occurs in particle physics.

Symmetry breaking describes how symmetrical systems can suddenly show a preference for one direction over another. As a simple example, imagine balancing a pencil on its tip. Viewed from the top, the balanced pencil appears symmetrical, but after a time it will fall and point in a single direction.

This concept applies to many physical systems, including superconductors, but it was Nambu who extended the theory to fundamental particles, according to John Ellis, a theorist at CERN, Europe's particle-physics laboratory near Geneva, Switzerland. Physicists now think that symmetry breaking is responsible for the creation of the Higgs boson, the particle that is believed to endow all other particles with mass,

and which is a target of CERN's Large Hadron Collider.

Kobayashi and Maskawa showed how violation of symmetry could create more matter than antimatter in the Universe — a long-standing problem in particle physics. In the early 1970s, the pair showed that the interactions of quarks via a fundamental force, called 'the weak force', could cause 'CP-violation', a phenomenon by which some particles decay in a different way from their anti-matter counterparts.

"They wrote down this huge expression whose physical interpretation is the violation of symmetry between matter and antimatter," says Ken Peach, a physicist at the University of Oxford, UK. The equations also predicted a third family of quarks (the particles that make up protons and neutrons in an atom's nucleus).

The idea of a thirds family seemed "far-fetched" at the time, says Ellis, but Kobayashi and Maskawa's work has since been verified by two high-energy experiments. The Belle experiment at KEK and the BaBar experiment at the Stanford Linear Accelerator Center in California both measured the decay of particles made of bottom quarks. The physicists' predictions were borne out to a high degree of accuracy. ■

Alison Abbott and Geoff Brumfiel

To read coverage of the Nobel Prize in Chemistry, which had not been announced as *Nature* went to press, visit www.nature.com.

Still, Walport says that the Wellcome Trust is not anticipating any need to alter its spending. Like most charities, it gives only a small fraction of its endowment to research each year — around £500 million in 2007. That relatively low rate of spending combined with careful management ahead should allow the charity to continue normal operations. "We are weathering the storm as well as can be expected," Walport says.

Cold Spring Harbor's US\$300 million endowment is also down, says Stillman, by an estimated 5% as of 31 July. And the laboratory, which supports a broad range of basic biological science, depends on additional philanthropic and corporate support to fund its roughly \$120-million operating budget. Stillman says that the money often comes either directly from Wall Street firms or from wealthy investors. The latest fundraising efforts are on track, he says. But "obviously the situation will affect our income down the road".

Rooney says that corporations may be the first to cut back on their giving. "Corporate donations are largely driven by changes in profitability," he says.

Lean times could be especially difficult for charities that depend entirely on annual donations to fund research. "Donors may raise their bar," warns Deborah Brooks, co-founder of the Michael J. Fox Foundation for Parkinson's Research in New York. Brooks says she thinks that wealthy individual donors will probably continue to give large sums but "their first instinct is to give to fewer people".

Stillman says he expects that Wall Street's woes will have a "significant" effect on research universities and charities across the state. Indeed, one important source of funding has dried up for good. Lehman Brothers, which filed for bankruptcy last month, donated roughly \$7 million to biomedical research in 2007. Most of the money went to hospitals and research centres in the New York area, with the largest amount, \$6 million, going to Weill Cornell Medical College's Lehman Brothers Lung Cancer Research Center. A spokesperson for the school declined to comment on the centre's future, saying only that it was "grateful for the many years of support by Lehman Brothers". ■

Geoff Brumfiel

Kavli invests in institutes

US philanthropist Fred Kavli has embarked on a second round of grants of up to \$5 million to each of the 15 research institutes that were established in his name. The grants, Kavli told *Nature*, are conditional on the institutes finding gifts from other donors so that they can establish endowments of \$20 million.

Kavli, who made his \$600-million fortune through real estate and selling his company — which had become a leader in supplying aeronautic and automotive sensors — says that the second round could take five years or more. "Right now the market hasn't been very kind to us," he says. "We are hedging a little on this."

Seven years ago, when Kavli established the first of the institutes, at the University of California at Santa Barbara, he developed what became a relatively strict formula:



Fred Kavli is hoping to support 20 institutes.

on average each university got \$7.5 million, to be supplemented and used however — for a building, for operating expenses, or for endowed professorships.

Kavli got the name. John Carlstrom, director of the Kavli

Institute for Cosmological Physics at the University of Chicago in Illinois, says that his university was not the only one that thought the gift was relatively small to be exchanged for the naming rights to an entire institute. But many assumed, apparently correctly, that more money would be coming later.

The first of the second round of gifts was completed in May to the Kavli Institute for Particle Astrophysics & Cosmology to Stanford University in Menlo Park, California. A deal with the University of California at Santa Barbara is in the works as fundraisers find matching gifts.

Kavli, 81, says that he views the institutes, which specialize in nanoscience, neuroscience, astrophysics and theoretical physics, as his legacy — and as more important than the three \$1 million eponymous prizes that were awarded for the first time this year. But he says that he is not done yet; he eventually envisions 20 institutes. ■

Eric Hand

Vapour spies to reveal climate clues

Scientists will begin collecting the first near-real-time measurements of the isotopes in water vapour on Hawaii's Mauna Loa this week, trialling what could prove to be a new way to study climate and weather systems.

The month-long experiment will deploy a trio of laser-based instruments to measure the isotopic composition of water vapour at least once every few minutes. That information can then be used to create a kind of 'life history' of the vapour, whether the source is a nearby forest, evaporation from the ocean or a cold air mass descending from higher in the atmosphere.

"This could be a new way of detecting fundamental changes in Earth's atmospheric circulation," says Joe Galewsky, an atmospheric scientist at the University of New Mexico in Albuquerque.

He says that Mauna Loa was the obvious place to conduct the experiment because of its sheer elevation and its location in the subtropics, which is influenced by both tropical and northerly atmospheric currents. These same attributes made the volcano an ideal place for the atmospheric carbon-dioxide counts recorded by the late Charles David Keeling, a long-time affiliate of the Scripps Institution of Oceanography in San Diego, California.

At almost 3,400 metres above sea level, the main observatory on Mauna Loa is comfortably above the fray of surface air currents. Galewsky and his team conducted a trial-run on nearby

Mauna Kea using flask samples in 2006 and found that the air currents seem to be governed by large-scale atmospheric processes rather than local weather trends (J. Galewsky *et al. Geophys. Res. Lett.* doi:10.1029/2007GL031330; 2007).

The research ultimately relies on minute differences in the way the isotopologues of water are affected by variables such as evaporation, transpiration and condensation. Heavy water (in which the molecule has an extra neutron) precipitates out of the atmosphere faster than the more common 'light' water isotopologue, and light water evaporates faster.

Scientists commonly study isotopic ratios in things such as ice cores and ocean sediments, but extracting them from water vapour in the air is more difficult. The advent of lasers that can be tuned to the specific absorption frequencies of individual isotopes made this work easier, as did the fact that such devices are now portable.

The team plans to test three such devices simultaneously: one from NASA's Jet Propulsion Laboratory in Pasadena and two commercially available devices that use slightly different technologies. The team will also gather air samples for comparison back in the lab. "We really have to convince ourselves that this works," says David Noone, a climatologist at the University of Colorado at Boulder, who is working on the project.

The next — and for many the most important — step is tying these measurements back to



Streams of vapour above Mauna Loa, Hawaii.

satellite data. A comparison of this sort was first published last year by Noone and colleagues in *Nature* (J. Worden *et al. Nature* 445, 528–532; 2007). In that study, the team analysed water-vapour isotopes using measurements from a passive emissions sensor aboard NASA's Aura satellite, one of two that will be collecting data during the current project.

"This is brand new. This is something that really revolutionizes the study of water vapour," says Gavin Schmidt, a climate modeller at the

Slime and fleas feature in Ig Nobel awards

Slime moulds exhibit the kind of "contemplative behaviour" that Hamlet is famous for, muses Toshiyuki Nakagaki of Hokkaido University in Japan. "Hamlet couldn't decide what to do so he did nothing. Then he chose a strong course of action." Similarly, slime moulds stop moving when they encounter a toxic substance. "But once action occurs, it too is strong," Nakagaki says.

The slime mould's puzzle-solving ability — Shakespearean or otherwise — is a discovery that is unlikely to change the world, but it won Nakagaki and his colleagues an Ig Nobel Prize for cognitive science last week at the annual event held at Harvard University in Cambridge, Massachusetts. Their research, published in

Nature (T. Nakagaki, H. Yamada & Á. Tóth *Nature* 407, 470; 2000), showed that slime moulds looking for food have "the ability to find the minimum-length solution between two points in a labyrinth".

Subsequently, the team has found that moulds can find the shortest path between 30–50 points, which is something even supercomputers cannot yet work out. "We can't even check the mould's solution," notes Nakagaki, "but it looks good."

From the cognitive feats

of slime mould to the physical prowess of fleas: Marie-Christine Cadiergues of the National Veterinary School in Toulouse, France, and her colleagues won the biology Ig Nobel Prize for showing that dog fleas can jump 2 centimetres

higher and 10 centimetres farther than cat fleas. "In the daily life of a vet, it's a good thing to know," Cadiergues says.

Other recipients of this year's prizes for amusing research included psychologists Charles Spence of the University of Oxford, UK, and Massimiliano Zampini of the University of Trento in Italy, who won the Nutrition Ig Nobel for research involving eating Pringles potato chips (crisps).



D. SCHARF/SCIENCE PHOTO LIBRARY



MUSE

The dripping taps of climate change.

www.nature.com/news

WWW.JUPITERIMAGES.COM



F. LANTING/CORBIS

NASA Goddard Institute for Space Studies in New York, who is not involved in the study. Schmidt calls water isotopes “the most super-duper fantastic thing ever”, and says that these types of data could fill an enormous gap in the scientific record by allowing scientists to analyse where water vapour comes from, what it is doing and where it is going.

“The data that are going to come out of this experiment really will allow us to go forward with confidence and know what it is we are modelling,” Schmidt says. ■

Jeff Tollefson

The pair showed that when the sound of chomping on a Pringle is amplified, people believe it is fresher than it really is.

The peace prize went to Klaus Peter Rippe, who chairs the Swiss Federal Ethics Committee on Non-Human Biotechnology, which is championing the idea that all living beings, including plants, have dignity (see *Nature* 452, 919; 2008). “Most people in Switzerland have no problem with this concept, though we’re aware that many people elsewhere find our ideas ridiculous,” he says.

Ironically, Rippe says, the 1952 Nobel Peace Prize was won by German-French physician Albert Schweitzer for, in part, promoting “reverence for life. Fifty years later, we get the Ig Nobel instead.” ■

Steve Nadis

Entire-paper plagiarism caught by software

When Eric Le Bourg, a French biogerontologist, came across a paper in a Korean journal recently, he almost fell off his chair; the entire article — text and graphs included — had been taken from one of his earlier articles. “It was plagiarism from beginning to end,” he says. “I was astonished; it was pure cut and paste.”

Such blatant copying of an entire article is not unknown, says Harold Garner, a researcher at the University of Texas Southwestern Medical Center in Dallas. Garner’s team has used its eTBLAST text-matching software to build *Deja Vu*, a continually updated database that already holds some 75,000 abstracts listed in Medline that seem highly similar. His team has so far found dozens of near-100% clone papers.

Garner estimates that among the 181 papers they have identified so far as duplicates, 85% of the text is similar on average, but one-quarter share close to 100%. For a full list of the most similar pairs of articles see <http://tinyurl.com/52s5e3>. There are currently 22 ‘repeat offenders’ in the database. These are authors who have published at least two articles that do not share authors (and so are putative or known plagiarisms). On average these people have ‘authored’ four papers, ranging from two to ten, and spanning 12 countries.

Le Bourg’s paper, “A review of the effects of microgravity and hypergravity on aging and longevity” was published in the Elsevier journal, *Experimental Gerontology* (E. Le Bourg *Exp. Gerontol.* 34, 319–336; 1999). The duplicate, by Hak-Ryul Kim, who listed his affiliation as the biology department of Korea University in Seoul, was published a year later in the *Korean Journal of Biological Sciences* (H.-R. Kim *Kor. J. Biol. Sci.* 4, 231–237; 2008).

Le Bourg and the editors of *Experimental Gerontology* have tried to investigate further, but to no avail. They contacted authorities at Korea University but got no response, Le Bourg says. E-mails to Kim were not returned. He seems to have left the university, says Bourg, who hasn’t been able to track down Kim’s current affiliation. Meanwhile, the *Korean Journal of Biological Sciences* has ceased publication.

With the trail gone cold, *Experimental Gerontology* intends to publish a note in its next issue stating that its editors have done their best to elucidate the case, and that “in the absence of any explanation, we believe that this is plagiarism of our article that we want to bring to the attention of the scientific community”.

Garner has begun to systematically contact editors and authors of the duplicates he has identified to assess how other cases have been followed up, and is submitting the results for publication. Many journal editors seem reluctant to pursue cases of plagiarism, and half of the articles that editors are alerted to remain uncorrected, Garner says. Few journals have communicated their retraction decision to PubMed, the most widely used abstracts database.

But the wider availability of tools to detect duplicated text is empowering editors. John Loadman, an editor of *Anaesthesia and Intensive Care*, who is a researcher at the Royal Prince Alfred Hospital near Sydney, Australia, is one of several editors who been

“I was astonished; it was pure cut and paste.”

using eTBLAST. He is ‘policing’ the anaesthesia literature and says that he has already found three

cases of duplication. Other publishers are using an anti-plagiarism tool called CrossCheck, which employs text-matching algorithms by iParadigms, a software company based in Oakland, California.

Many of the duplicates in *Deja Vu* come from non-English-speaking countries, and some scientists have asserted that a degree of plagiarism is justified as a way of improving the English of their texts (see *Nature* 449, 658; 2007). “There definitely is a cultural component,” says Garner, “but this appears to be an equal-opportunity behaviour, with scientists from across the world involved.”

When confronted with their plagiarism, some researchers can be brazen. One offender, whose paper shared 99% of its text with an earlier report, wrote to Garner: “I seize the opportunity to congratulate [the authors of the original paper] for their previous and fundamental paper — in fact that article inspired our work.” ■

Declan Butler



MUSE

The dripping taps of climate change.

www.nature.com/news

WWW.JUPITERIMAGES.COM



F. LANTING/CORBIS

NASA Goddard Institute for Space Studies in New York, who is not involved in the study. Schmidt calls water isotopes “the most super-duper fantastic thing ever”, and says that these types of data could fill an enormous gap in the scientific record by allowing scientists to analyse where water vapour comes from, what it is doing and where it is going.

“The data that are going to come out of this experiment really will allow us to go forward with confidence and know what it is we are modelling,” Schmidt says. ■

Jeff Tollefson

The pair showed that when the sound of chomping on a Pringle is amplified, people believe it is fresher than it really is.

The peace prize went to Klaus Peter Rippe, who chairs the Swiss Federal Ethics Committee on Non-Human Biotechnology, which is championing the idea that all living beings, including plants, have dignity (see *Nature* 452, 919; 2008). “Most people in Switzerland have no problem with this concept, though we’re aware that many people elsewhere find our ideas ridiculous,” he says.

Ironically, Rippe says, the 1952 Nobel Peace Prize was won by German-French physician Albert Schweitzer for, in part, promoting “reverence for life. Fifty years later, we get the Ig Nobel instead.” ■

Steve Nadis

Entire-paper plagiarism caught by software

When Eric Le Bourg, a French biogerontologist, came across a paper in a Korean journal recently, he almost fell off his chair; the entire article — text and graphs included — had been taken from one of his earlier articles. “It was plagiarism from beginning to end,” he says. “I was astonished; it was pure cut and paste.”

Such blatant copying of an entire article is not unknown, says Harold Garner, a researcher at the University of Texas Southwestern Medical Center in Dallas. Garner’s team has used its eTBLAST text-matching software to build *Deja Vu*, a continually updated database that already holds some 75,000 abstracts listed in Medline that seem highly similar. His team has so far found dozens of near-100% clone papers.

Garner estimates that among the 181 papers they have identified so far as duplicates, 85% of the text is similar on average, but one-quarter share close to 100%. For a full list of the most similar pairs of articles see <http://tinyurl.com/52s5e3>. There are currently 22 ‘repeat offenders’ in the database. These are authors who have published at least two articles that do not share authors (and so are putative or known plagiarisms). On average these people have ‘authored’ four papers, ranging from two to ten, and spanning 12 countries.

Le Bourg’s paper, “A review of the effects of microgravity and hypergravity on aging and longevity” was published in the Elsevier journal, *Experimental Gerontology* (E. Le Bourg *Exp. Gerontol.* 34, 319–336; 1999). The duplicate, by Hak-Ryul Kim, who listed his affiliation as the biology department of Korea University in Seoul, was published a year later in the *Korean Journal of Biological Sciences* (H.-R. Kim *Kor. J. Biol. Sci.* 4, 231–237; 2008).

Le Bourg and the editors of *Experimental Gerontology* have tried to investigate further, but to no avail. They contacted authorities at Korea University but got no response, Le Bourg says. E-mails to Kim were not returned. He seems to have left the university, says Bourg, who hasn’t been able to track down Kim’s current affiliation. Meanwhile, the *Korean Journal of Biological Sciences* has ceased publication.

With the trail gone cold, *Experimental Gerontology* intends to publish a note in its next issue stating that its editors have done their best to elucidate the case, and that “in the absence of any explanation, we believe that this is plagiarism of our article that we want to bring to the attention of the scientific community”.

Garner has begun to systematically contact editors and authors of the duplicates he has identified to assess how other cases have been followed up, and is submitting the results for publication. Many journal editors seem reluctant to pursue cases of plagiarism, and half of the articles that editors are alerted to remain uncorrected, Garner says. Few journals have communicated their retraction decision to PubMed, the most widely used abstracts database.

But the wider availability of tools to detect duplicated text is empowering editors. John Loadman, an editor of *Anaesthesia and Intensive Care*, who is a researcher at the Royal Prince Alfred Hospital near Sydney, Australia, is one of several editors who been using eTBLAST. He is ‘policing’ the anaesthesia literature and says that he has already found three

“I was astonished; it was pure cut and paste.”

cases of duplication. Other publishers are using an anti-plagiarism tool called CrossCheck, which employs text-matching algorithms by iParadigms, a software company based in Oakland, California.

Many of the duplicates in *Deja Vu* come from non-English-speaking countries, and some scientists have asserted that a degree of plagiarism is justified as a way of improving the English of their texts (see *Nature* 449, 658; 2007). “There definitely is a cultural component,” says Garner, “but this appears to be an equal-opportunity behaviour, with scientists from across the world involved.”

When confronted with their plagiarism, some researchers can be brazen. One offender, whose paper shared 99% of its text with an earlier report, wrote to Garner: “I seize the opportunity to congratulate [the authors of the original paper] for their previous and fundamental paper — in fact that article inspired our work.” ■

Declan Butler

Change of guard for Australian science

Astronomer Penny Sackett is to be Australia's next chief scientist, taking over from Jim Peacock in November, the government announced last week.

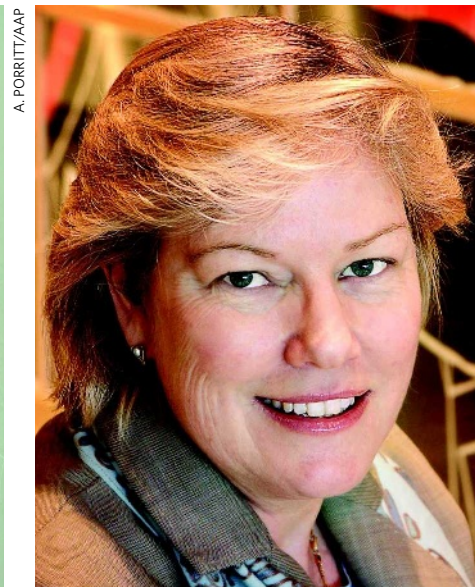
It is not the only change at the top of Australian science: geologist Megan Clark will become head of the national research agency, the CSIRO, in January. Meanwhile, the Australian Stem Cell Centre is without a permanent leader after its chief executive was sacked in July, and the entire board resigned amid controversy in September. An interim chair of the board, Graham Macdonald, has been appointed.

Sackett's appointment has been greeted with broad approval by researchers. The US-born scientist, originally from Lincoln, Nebraska, has lived in Australia since 2002, when she was recruited to head the Mount Stromlo and Siding Spring Observatories, and to run the astronomy and astrophysics department at the Australian National University in Canberra. Sackett has also been involved in the governance of the Gemini Observatory, the Hubble Space Telescope and the Giant Magellan Telescope.

The job of chief scientist will be made a full-time post, upgraded from part-time, once Peacock, a CSIRO plant scientist, leaves next month. The purpose of the role is to provide the government with independent advice on scientific matters relating to policy. Sackett will need to shake off long-standing concern about political interference in the post. Her predecessors have been criticized for siding with the government on some controversial decisions, such as Peacock's support of an environmentally controversial pulp mill and his predecessor Robin Batterham's dismissal of the Kyoto Protocol on climate change (see *Nature* 435, 398; 2005).

Sackett says she discussed these issues openly with the government before accepting the post. "The ground rules we agreed on are complete independence of advice and a 'no-surprises' policy. It would not be productive if the first time the government heard the advice was by reading it in the newspaper," she says. "The advice would be independent and not adversarial," Sackett adds.

Some are concerned about the fact that the chief scientist reports directly to the minister for innovation, industry, science and research, rather than the prime minister. "It is very important that the chief scientist is representative of all Australian science, and not just the



New guard: Penny Sackett (left) and Megan Clark will take two of Australia's top science jobs.

science within the ministry of innovation," says Kurt Lambeck, president of the Australian Academy of Science. But Sackett is confident that there are other lines of communication, including the Prime Minister's Science, Engineering and Innovation Council.

When Kevin Rudd's Labor government came to power at the end of 2007, it instigated a swag of reviews, and their findings are beginning to emerge. One such review, released last month, calls for university research to be fully publicly funded — currently, universities must find some of their costs from other sources. The review

also calls on the government to step up research spending to match other nations in the Organisation for Economic Co-operation and Development. "The chief scientist will play a major role in shaping the government's response to

these reviews, so she has got a real challenge immediately on her plate," says Lambeck.

This is the first time the post has been held by a woman — a novelty shared by Clark, who will be the first female head of the CSIRO. She too faces a challenging time when she leaves her job as a senior executive for global mining giant BHP Billiton. She takes the helm of a battered organization that has suffered a decade of funding cuts, restructuring, job losses and diminishing staff morale. Lambeck is optimistic that she can meet the demands. "She is an independent person with a lot of

experience, and I think she will do what she sees as best for the organization," he says.

Things are not looking as rosy at the Australian Stem Cell Centre in Melbourne, a biotech founded with public money in 2003 by Alan Trounson, who now heads the California Institute for Regenerative Medicine in San Francisco. On 5 September, the board resigned after accusing stakeholder universities of not supporting its independence. A week earlier, the board had sacked the centre's chief executive, Stephen Livesey, because of differences over the commercial direction of the centre. During subsequent consultations on the centre's future direction, the board and some university partners clashed over control of the board.

"We saw resignation as the only option available," says Vicki Sara, ex-chair of the centre's board and chancellor of the University of Technology Sydney. Sara says that, by moving aside, the outgoing board hopes that a fresh approach will attract long-term government funding.

Original expectations that the centre would be financially self-sufficient after current government funding expires in June 2011 were "unrealistic," says interim chair Macdonald. The immediate objective is to redesign the centre's business plan. A minimal model, says Macdonald, could provide the stem-cell community with management, commercialization advice and technology services on intellectual property. "The objective is to resolve those questions within the next month and a half," he says. ■

Carina Dennis

"The ground rules we agreed on are complete independence of advice and a 'no-surprises' policy."

Q&A

The future for Howard Hughes



B. RIES/HHMI

Robert Tjian is a biochemist and molecular biologist at the University of California, Berkeley, who has made significant discoveries about DNA transcription factors. He co-founded the biotechnology company Tularik, which was sold in 2003 for US\$1.3 billion. On 1 April next year, Tjian will take over as president of the privately funded Howard Hughes Medical Institute (HHMI) in Chevy Chase, Maryland. He spoke to *Nature* about his plans.

What are your priorities for the HHMI?

The main job is to keep identifying the best people out there who are doing the kind of science that we like — creative, original, somewhat risky — and to give them enough resources to follow their dreams.

Do you want to expand the HHMI's activities in any field?

Its present direction is already very promising: for example, getting more into chemical engineering, chemistry and bioengineering. We might also look into plant biology and biofuels.

Will the economic crisis affect the HHMI?

Since 2001, when we had the last downturn, Hughes has become much more adept at dealing with these downturns and upturns. My impression from the board is that their endowment is probably just going to be flat, but not negative.

What is your arrangement with Berkeley?

I'll continue to have a lab there, although it'll be pared down. A small portion of my lab may be moved to Janelia Farm on the HHMI campus. But I'm not going to be doing anything with

biotech, because the president of the HHMI doesn't do any consulting for companies.

Is Janelia Farm living up to its promise as an interdisciplinary research centre?

Janelia was a worthy experiment. We don't know how successful it'll be yet, but it has had a good start and needs time to mature.

Will you try other, similar experiments?

It would be nice if we could do something of that magnitude. I'm optimistic that the economy will eventually improve and our endowment will grow so that more unusual, creative programmes could be discussed.

What are your plans for the HHMI's education and international programmes?

The international programme is terrific in terms of bang for its buck. I will certainly see if we can expand it. Tom [Cech, current HHMI president] went to the HHMI with a very strong vision about education, particularly at the lower levels. I think that's a really difficult nut to crack. I'm more inclined to look at the graduate and upper-division undergraduate opportunities in this country, where I still feel the American education system really shines.

What might this entail?

For example, technology in the laboratory is changing rapidly, so it would be nice to expose our students to the modern methodology instead of teaching them the same labs we've been teaching them for the past hundred years. Another thing I'd like to do — even though I am a hardcore, reductionist biochemist — is to connect the molecular aspects of biology to really tough problems such as disease. I'm talking about teaching our students to think molecularly but to apply it to human conditions.

What is your management style?

I enjoy working with people who are fun to work with and who are as intense as I am. I'm still working seven days a week because I like it. I don't consider anything I do as work. I tell my students, I'm not expecting you to come to lab because you're doing work — you come to lab because you're having fun. I come in at weekends not because I think it's an obligation to my work, but because I genuinely enjoy what I do.

What excites you about biology today?

We're in an epoch of the most rapid growth in the biological sciences: genome sequencing methodologies moving at the speed of light; microfluidics; microscopy advancing at amazing rates; mass spectrometry; systems biology; and computational tools. And there are so many interesting, challenging problems — diagnostics, therapeutics, biofuels, you name it.

Interview by Erika Check Hayden.

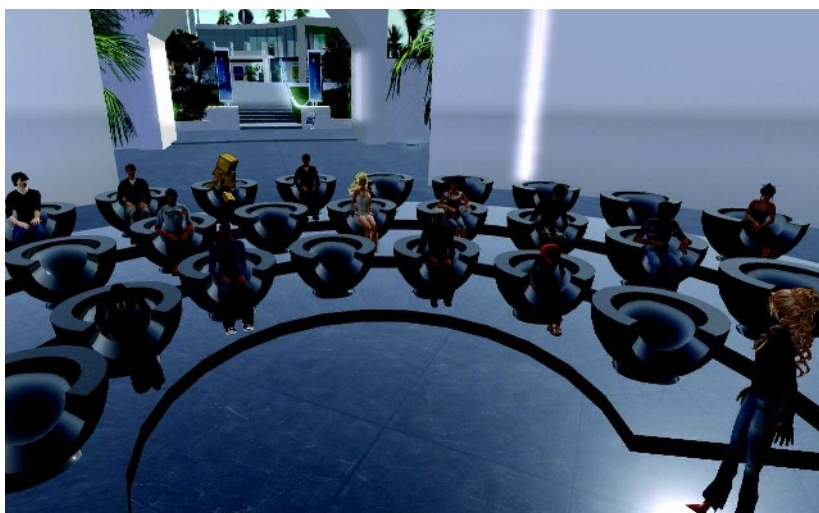
SNAPSHOT

The *in silico* show

A motley crew of curvaceous women, cardboard men and other avatars gathered last Thursday on *Nature's* island in Second Life. The occasion for this virtual world get-together was the premiere of *Missions in Space-Time*, a five-part documentary series produced by *Nature Video*.

The films capture conversations between young researchers and Nobel prizewinning physicists at the 2008 Nobel Laureate Meeting on Lindau Island in Germany. The first instalment discusses hopes and dreams for the Large Hadron Collider — the particle accelerator at CERN, Europe's particle-physics laboratory near Geneva in Switzerland.

Part one is available for download now; subsequent instalments will be released weekly via iTunes and online at www.nature.com/video/lindau.



UK launches department to focus on climate change

UK prime minister Gordon Brown has created a Department of Energy and Climate Change.

The announcement, made on 3 October as part of a wider reorganization of the government, was widely praised by environmentalists, who hope that the department will unify the nation's efforts to reduce greenhouse-gas emissions. It will be led by Ed Miliband, who is regarded as a close ally of the prime minister and whose brother, David, is the foreign secretary.

Previously, climate change and energy policy in the government were spread across several agencies. "It was a bit of a mess, actually," says Ben Caldecott, head of the environment unit at Policy Exchange, a London-based think tank. Caldecott calls the new climate-change department "a step in the right direction", but says more must be done to unify policy.

The government reshuffle has also ushered in a new science minister. Paul Drayson, a multimillionaire and former head of Defence Equipment and Support, the UK defence ministry's procurement organization, will replace Ian Pearson. Drayson holds a PhD in robotics from Aston University.

US academic fleet cuts operations as budget bites

Budget freezes and high fuel costs will force one vessel in the US academic oceanographic fleet — probably from an East Coast institution — to remain docked throughout 2009. Two other members of the fleet's 23 ships will run on only partial schedules.

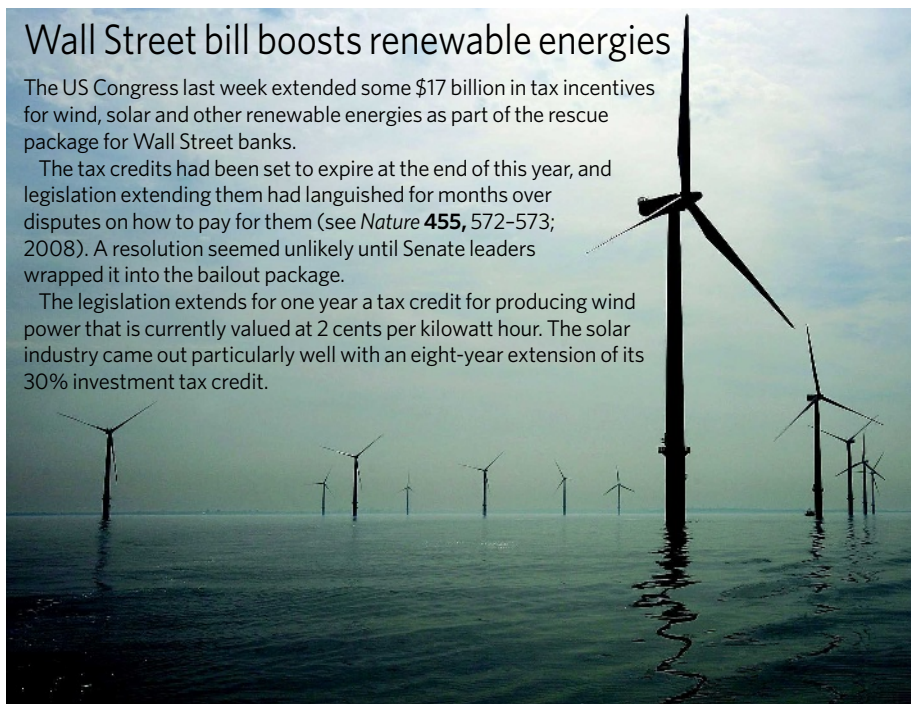
At their annual meeting in Arlington, Virginia, last week, council members of

Wall Street bill boosts renewable energies

The US Congress last week extended some \$17 billion in tax incentives for wind, solar and other renewable energies as part of the rescue package for Wall Street banks.

The tax credits had been set to expire at the end of this year, and legislation extending them had languished for months over disputes on how to pay for them (see *Nature* 455, 572–573; 2008). A resolution seemed unlikely until Senate leaders wrapped it into the bailout package.

The legislation extends for one year a tax credit for producing wind power that is currently valued at 2 cents per kilowatt hour. The solar industry came out particularly well with an eight-year extension of its 30% investment tax credit.



C. FURLONG/GETTY IMAGES

the University-National Oceanographic Laboratories System (UNOLS) said the cutbacks were necessary to save money. Despite sidelining ships, UNOLS still faces one of its largest ever deficits, of \$19 million, in its \$100-million budget to fund next year's fleet.

Final plans to select the ships and strategies to cover the deficit will be made by the end of October.

Italy to create biomedical-research funding system

The Italian government has pledged to create a formal system for the competitive distribution of funds for biomedical research, which will conform to international standards of peer review and transparency.

The statement of intent has partially mollified biologists, some of whom last year protested the murky allocation of research funds by the ISS, Italy's health-research institute in Rome. Stem-cell researchers were particularly angry about how funds earmarked for their field were distributed (see *Nature* 450, 320; 2007).

The health ministry, which made the announcement on 1 October, has not yet indicated when the system would come into effect.

Eli Lilly wins race to take control of ImClone

Pharmaceuticals giant Eli Lilly, of Indianapolis, Indiana, seems to have beaten its rival Bristol-Myers Squibb in the

competition for ImClone, a biotechnology firm based in New York. On 6 October, Eli Lilly and ImClone announced that they had agreed to a US\$6.5-billion merger.

Bristol-Myers Squibb, also of New York, owns 17% of ImClone and has an agreement to market its blockbuster anticancer drug Erbitux (cetuximab), a monoclonal antibody. In July Bristol-Myers Squibb failed in a bid to buy the biotech company, offering \$60 a share, when ImClone's shares were trading at around \$40–45. A follow-up offer of \$62 a share was also turned down. The Eli Lilly deal values the company at \$70 a share.

Bristol-Myers Squibb is unlikely to bid again, so agreement with Eli Lilly is being seen as the endgame of the protracted wrangling over the company, Erbitux and promising follow-up monoclonal antibodies.

United States agrees to sell nuclear fuel to India

India can buy nuclear fuel and technologies from the United States, now that the US Senate has approved the '123 deal'.

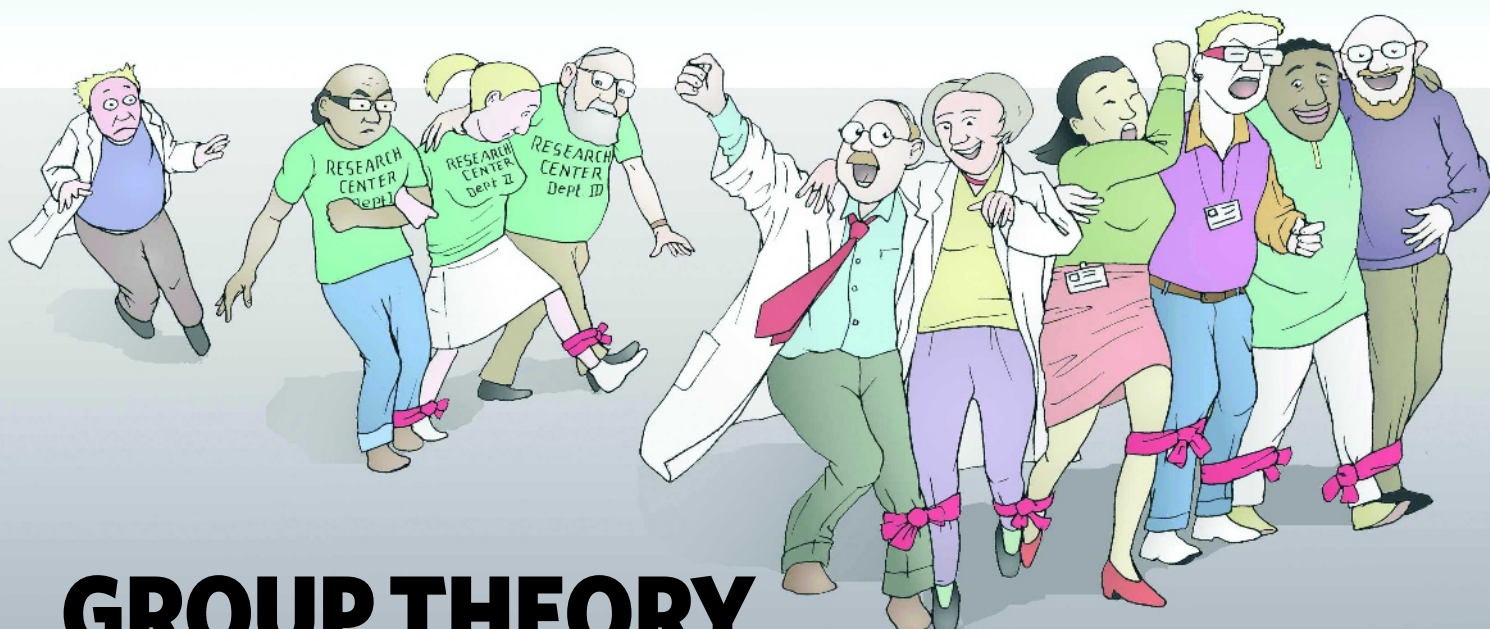
The vote, along with earlier waivers from the Nuclear Suppliers Group and the International Atomic Energy Agency, makes India the only country with nuclear weapons that can legally engage in nuclear commerce without first having signed the Nuclear Non-Proliferation Treaty.

US secretary of state Condoleezza Rice was to have sealed the deal on a visit to New Delhi last weekend. As *Nature* went to press, President George W. Bush was expected to sign the deal on 8 October.



Some of the US research fleet may be suspended.

SOEST/UNIV. HAWAII



GROUP THEORY

What makes a successful team? **John Whitfield** looks at research that uses massive online databases and network analysis to come up with some rules of thumb for productive collaborations.

Flip through any recent issue of *Nature*, including this one, and the story is there in black and white: almost all original research papers have multiple authors. So far this year, in fact, *Nature* has published only six single-author papers, out of a total of some 700 reports. And the proportions would be much the same in any other leading research journal.

Of course, there is nothing new about this: the scholars who study the folkways of science have been tracking the decline of the single-author paper for decades now. And they have followed the parallel growth of 'invisible colleges' of researchers who are separated by geography yet united in interest. But what is new is how their studies have been turbocharged by the availability of online databases containing millions of papers, as well as analytical tools from network science — the discipline that maps the structure and dynamics of all kinds of interlinked systems, from food webs to websites.

The result is a clearer picture of science's increasingly collaborative nature, and of the factors that determine a team's success. Funding agencies are not using this work to decide where the money goes — yet. But the researchers behind the analyses are willing to give tentative tips on what their work reveals. They also think that their studies point to rules of thumb that apply very broadly, whether you're looking for a gene or putting on a show.

The first question a researcher might ask him- or herself is: should I collaborate at all?

Perhaps the rarity of single-author papers would translate into higher impact? To answer this question, sociologist Brian Uzzi of Northwestern University in Evanston, Illinois, and his colleagues analysed more than 2 million patents¹, along with nearly 20 million papers published since 1955. They found that in the early 1950s, the most cited paper in any year was more likely to have been written by a single author than a team, but this pattern reversed decades ago. And the citation gap continues to widen.

"The image of the scientist alone at the workbench, plucking ideas out of the ether was true up to about the end of the Second World War," says Uzzi, "but not any more."

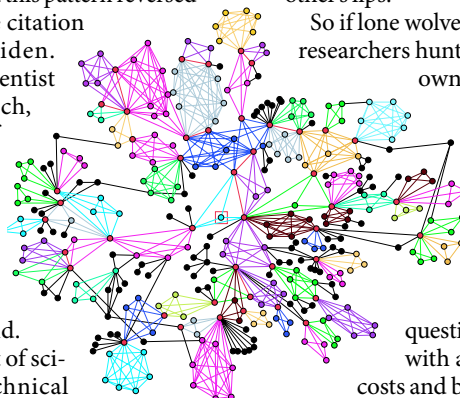
Uzzi doesn't know what drives this trend. It is not just a product of science's increasing technical complexity: the same pattern is seen in pencil-and-paper disciplines such as mathematics and the humanities. It is not just the Internet: author teams began to swell long before the online age, and the dawn of e-mail hardly affected that growth. And it is not just that large teams create many opportunities for future self-promotion: the pattern remains when self-citation is removed. Uzzi speculates that the increasing specialization of all fields plays a part, as may changing social

norms. Researchers have always swapped ideas and criticism, but when fields were small, authorship was not such an important mark of achievement. Reputation travelled by word of mouth, and everyone knew who had contributed the good ideas. Now, however, academia is too vast for that kind of informal credit assignment to work. So people need to get their ideas and their names into print, as well as on each other's lips.

So if lone wolves go hungry, who should researchers hunt with? Someone in their own discipline, or someone in another field? Should they build long-term relationships, or should they keep changing the people they work with?

Research is now revealing that these questions need to be answered with a careful weighing up of costs and benefits, rather than a list of absolute dos and don'ts: teams are most successful when they contain the right mix of specialism and diversity, and familiarity and freshness. And researchers are starting to find hints of how to strike this balance.

Uzzi and his team, for instance, looked at a sample of 4.2 million papers published between 1975 and 2005. Dividing universities into tiers based on the number of citations their researchers achieved, they found that teaming up with someone from another institution



J.H. VAN DIERENDONCK

REF. 3

N. MANDELL



Brian Uzzi has tracked changes in how citation rates relate to the number of authors on a paper.

of the same or higher tier reliably produced more highly cited work than teaming up with someone down the corridor.

"There's something about between-school collaboration that's associated with the production of better science," Uzzi told participants at a meeting of network scientists in Norwich, UK, in June. At the same meeting Pietro Panzarasa, from Queen Mary University of London, presented an analysis of 9,325 papers written by 8,360 authors submitted to the 2001 UK Research Assessment Exercise in business and management studies. He too found that between-institution collaborations had a higher average impact than did those within institutions.

Middle ground

As well as looking at where people worked, Panzarasa looked at how specialized they were. First he assigned researchers to disciplines by analysing the keywords in their papers, and then he measured each author's breadth of experience by looking at the fields of their co-workers. Social scientists are divided over whether specialization is the best strategy, he says. "It is beneficial for productivity and earning, but there is also evidence from banking and academia that being a generalist pays off." Panzarasa's data show that the most highly cited papers were written either by authors who worked mostly with others in their own field or by those who worked with people in a wide range of other disciplines. But between these peaks lay a trough: papers that had authors from an intermediate number of disciplines were the most poorly cited.

"Being extremely specialized allows you to exploit the benefits of being embedded in your discipline, such as reputation, con-

sensus building and controlling the flow of knowledge," says Panzarasa. "When you go to the other extreme you can take advantage of all the information coming from different pools of knowledge. But if you're somewhere in the middle, you have less success — unless you feel you can manage very high levels of interdisciplinarity, it might be better to stay in your discipline."

The most successful interdisciplinary authors, Panzarasa found, work with people who have independent authorship connections with each other, creating a tight social network. Panzarasa suspects that when these backup connections between colleagues are missing, the person in the middle can flounder as they try to process too many information streams. But his analysis also found that highly specialized workers who broaden their focus slightly produce more highly cited papers, as do those that exploit what social scientists call brokerage: bridging communication gaps between researchers who don't otherwise interact, and acting as a conduit for transferring knowledge from one field to another. Specialist brokers produced the most highly cited papers of any in his sample.

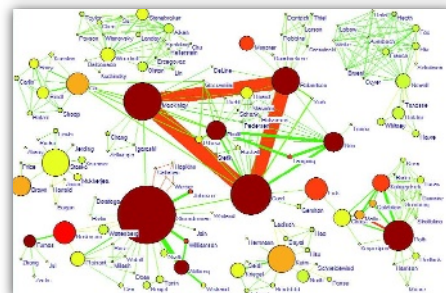
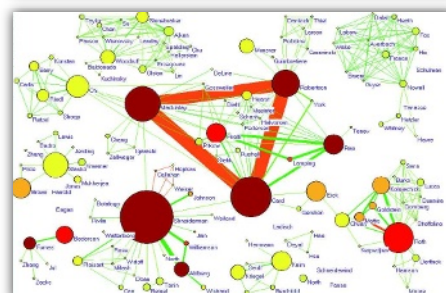
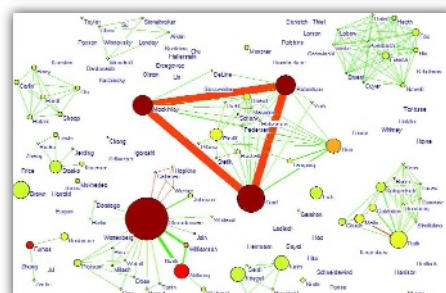
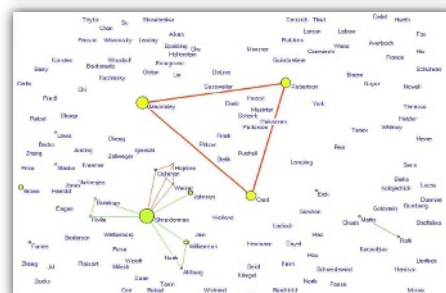
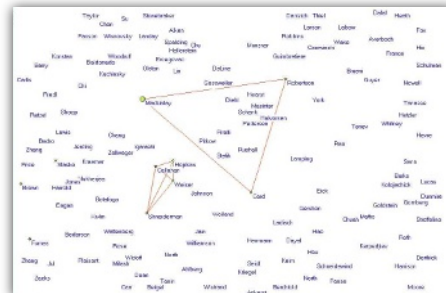
The lesson of these studies might seem to be that if you do decide to take the leap across disciplinary boundaries, then the more addresses and subjects you can cram onto an author list, the better. But not necessarily. All these surveys have looked for co-authorship patterns in the published literature, which means that they have a built-in bias: they look only at the collaborations that actually result in publication.

"The image of the scientist plucking ideas out of the ether was true up to about the end of the Second World War."

— Brian Uzzi

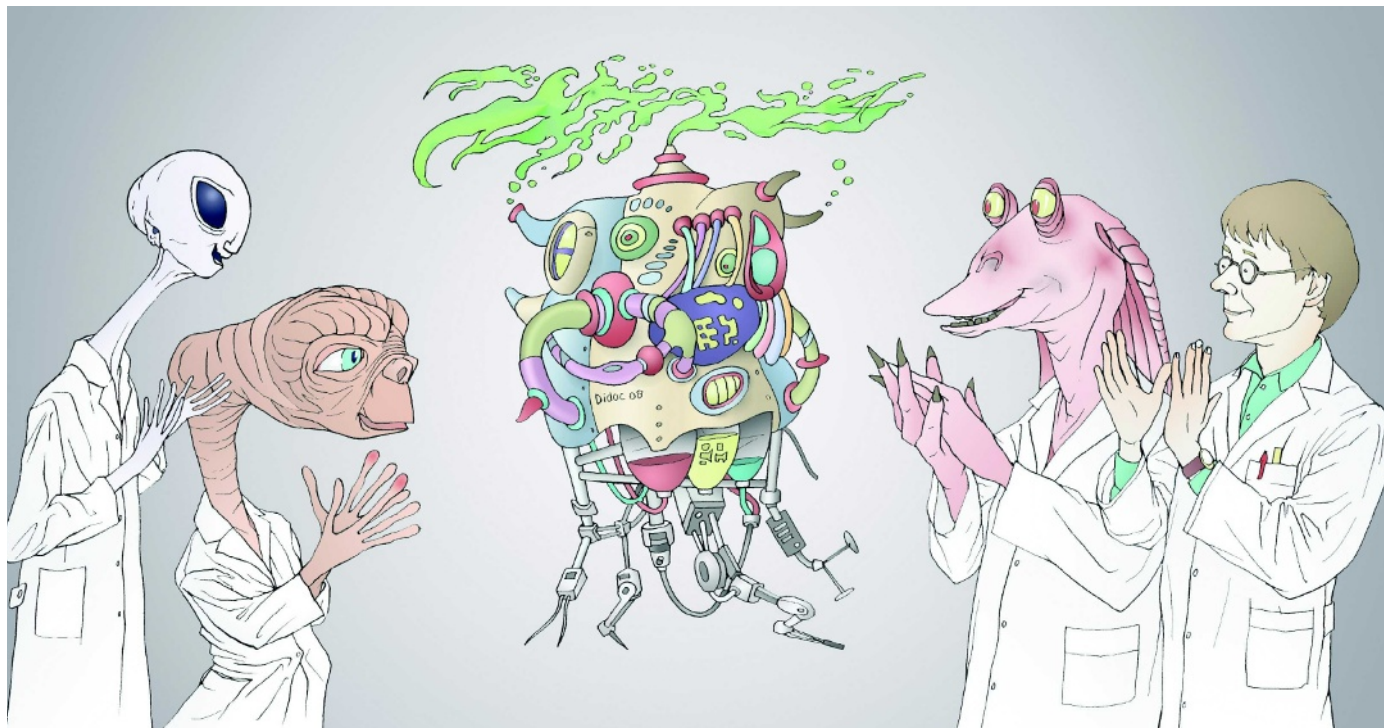
In fact, teams can also fail if they spread themselves too thinly. Jonathon Cummings, of Duke University's Fuqua School of Business in Durham, North Carolina, is monitoring more than 500 projects funded by the US National Science Foundation's information technology research programme, which creates cross-disciplinary teams of natural, social and computer scientists. He found that the most diverse teams were, on average, the least productive². "Projects that had more universities involved were at a greater risk of not publishing," says Cummings, as were those that covered multiple disciplines.

This apparent discrepancy is resolved by thinking of interdisciplinary research as a high-risk, high-reward business, explains Sam Arbesman, a mathematician at Harvard Medical School in Cambridge, Massachusetts, who has studied authorship networks. "A more diverse team isn't always better — it might be that you get more really good or



By mapping which authors collaborated with whom (lines) and when (colour of lines), Katy Börner and her colleagues show how networks extend with time and how the impact of both an author (thickness of nodes) and a partnership (thickness of lines) can grow.

W. K. L. VISVANATH & K. BÖRNER



J.H. VAN DIERENDONCK

really bad research,” he says. Still, there are ways to reduce the risks that the work won’t be publishable. Cummings found that if the principal investigators had a previous history of collaboration, their project was much more likely to be successful than if they had never written a paper together before. Such teams will have already paid the start-up costs of getting everyone familiar with one another’s approaches and languages; new teams should invest in travel and seminars, he says. “Familiarity adds a lot of value.”

Talent spotting

“We can spot projects that have been patched together at the last minute in response to the latest call for proposals,” says Suzanne Iacono, who directs the information technology research programme. “Reviewers say, ‘These people have never produced a paper before, and we’re going to give them \$15 million?’” The programme currently requires researchers to include plans for team-building in their proposals, but Iacono wants more than that. “I’d like to understand better the point at which bringing in more disciplines leads to a decline in knowledge production,” she says.

But it is a fine line between a collaboration that has found its groove and one that has fallen into a rut. And it’s not a line that people spot easily, because mature groups gravitate towards common ground and avoid areas of disagreement. Network scientists call this an echo chamber: a situation in which everyone tells everyone else what they want to hear, and a group that thinks it is performing well is really just mired in consensus.

To avoid stagnating, scientists think that teams need a stream of fresh input. And the optimum rate of turnover seems to depend on the size of the team. In a paper published

in *Nature* last year³, physicist Gergely Palla of the Hungarian Academy of Sciences in Budapest and his colleagues analysed networks of authorship on physics papers posted to the arXiv preprint server. They showed that teams with around 20 members had a better chance of surviving for a long period if they had a high rate of arrival and departure. For a team of three or four to persist, however, the opposite was true — they needed stability. Palla speculates that it’s easy to find two people you like well enough to form a long-term working relationship; in a big team, fall-outs are inevitable, but the whole can persist if the comings and goings are constant and low-level. Endurance is not the same as quality of output, of course, but, as Pallas says: “It’s hard to imagine that you would publish rubbish for a long time.”

But even small groups benefit from some turnover. Looking at a data set of nearly 90,000

papers published between 1955 and 2004 by 115,000 authors in 32 journals spread across the fields of social psychology, economics, ecology and astronomy, Luis Amaral, a network scientist at Northwestern, and his colleagues measured the proportion of authors who had worked with each other before⁴. Papers in high-impact journals showed a strikingly lower proportion of these repeated interactions than did papers in low-impact journals. “The patterns with repeat collaboration are very different and dramatic,” says Amaral. “In low-impact journals, people repeat collaborations almost all the time.”

When people choose collaborators, says Uzzi, who also worked on this analysis, they look for two opposing things: high-status individuals with a proven record and good resources, and newcomers who have lots of time and energy to devote to a project. The trick is to find the balance. “If you had to give people a rule of



L. TODD/DUKE PHOTOGRAPHY



A. CAMPBELL

Jonathon Cummings (left) and Luis Amaral study how a team’s composition affects its success.

thumb, you might want 60–70% of a team to be incumbents, and 50–60% repeat relationships,” Uzzi says. “That gets you into the bliss point across four very different scientific fields.”

And this is not just in science — the same, they found, goes for Broadway musicals. It typically takes six specialists to create and put on a musical: one each to write the music, lyrics and dialogue, plus a choreographer, director and producer. The most critically and financially successful musicals have an intermediate level of turnover within the creative team⁵. Amaral thinks there may be group properties that influence outcomes across all kinds of collective effort — “but we’ll need a lot of data to figure them out”, he says.

Uzzi has been mobbed by organizations seeking to locate their bliss points. “The president of a university called me up to ask how he can tell what areas he should be investing in,” he says. Corporations have been asking for tips on assembling work groups; venture capitalists want to know how to spot the next hot field; a delegation from the US National Institutes of Health (NIH) is interested in whether the work can help make funding decisions; and Uzzi has been invited to the offices of *Nature* and *Science*, as both journals strain after ways to detect the highest-impact papers.

First come, most cited?

Another issue is the opaque relationship between a paper’s citations and its science. A known trend is that the more a paper is cited, the more citations it attracts, which stretches small gaps in quality into chasms in citation count. The process can also reward novelty above merit — in a preprint posted online this September, physicist Mark Newman of the University of Michigan in Ann Arbor models and measures the effects of ‘first-mover advantage’ on citations, showing that it has no relation to the quality of the research. Those who are the first to publish in a new field are likely to garner more citations than those who publish later⁶. “Were we wearing our cynical hat today,” he writes, “we might say that the scientist who wants to become famous is better off writing a modest paper in next year’s hottest field than an outstanding paper in this year’s.”

There are also other networks to consider: analysing every paper published in the *Proceedings of the National Academy of Sciences* between 1982 and 2001, Katy Börner, who studies networks and information visualization at Indi-

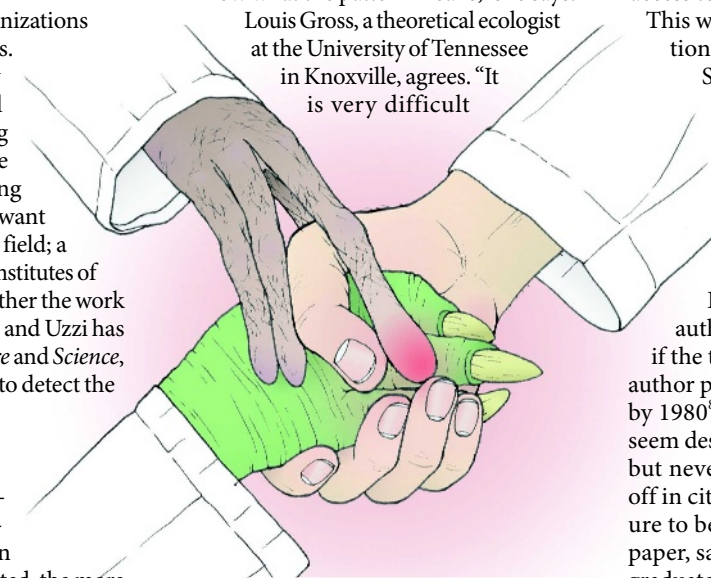
ana University in Bloomington, found that US authors are more likely to cite papers by workers at nearby institutions than from those on the other side of the country⁷. “People read widely,” she says, “but when it comes to filling the slot at the end of the paper, they also consider who they have to face again in the hallway or at the next conference.”

Such factors make some urge caution about using network analysis. At present, no one should be using such techniques to judge a collaboration’s likely performance, says Deborah Duran, head of the systemic assessments branch of the Office of Portfolio Analysis and Strategic Initiatives at the NIH. “We can see a pattern, but we don’t know what the pattern means,” she says.

Louis Gross, a theoretical ecologist at the University of Tennessee in Knoxville, agrees. “It is very difficult

“We can spot projects that have been patched together at the last minute.”

—Suzanne Iacono



to account for the effects of social networking in evaluating metrics of citations. Network analysis definitely has potential, but an awful lot of social science needs to be integrated with these analyses to ensure that they are applied in an equitable way,” he says. Gross has reviewed grant proposals for the European Commission; one risk in Europe, he says, is that if granting agencies place too much emphasis on encouraging international collaborations they will stunt development within institutions and nations.

But Duran does expect network studies to be an important part of what she calls “the emerging science of science management”. The NIH already uses data-mining tools devised by the company Discovery Logic, based in Rockville, Maryland, to see how grants connect to papers, citations, patents and products. Duran suggests that in the future, network analysis could be used to track the spread of new ideas, work out the best ways to disseminate information or to

target particularly well-connected individuals to work on emerging issues. “I think, hope and believe that this will become useful,” she says.

So, can a scientist looking to make the most of his or her talent really exploit these findings? Amaral says that network analysis might actually help young researchers to look beyond citation counts, which are dominated by a field’s obvious stars, and find other groups with a healthy mix of rookies and veterans and a productive rate of turnover. At present, a do-it-yourself approach would be difficult: mapping the networks and measuring scientific success requires access to subscription databases such as ISI and computing resources that are beyond the reach of the average web-surfing graduate student. But this is about to change: Börner and her colleagues are soon to release an open-access tool for analysing scholarly networks. This will allow researchers to map connections using free sources such as Google Scholar, as well as Indiana University’s database of 20 million publications, patents and grants and even its own bibliographic files.

But however finely honed scientists’ team-building strategies become, there will always be room for the solo effort. In 1963, Derek de Solla Price, the father of authorship-network studies, noted that if the trends of that time persisted, single-author papers in chemistry would be extinct by 1980⁸. In fact, many branches of science seem destined to get ever closer to that point but never reach it⁹. And whatever the payoff in citations might be, there’s still a pleasure to be had in seeing just your name on a paper, says Matt Friedman, a palaeontology graduate student at the University of Chicago and a member of *Nature*’s sextet of singleton authors¹⁰. “With any piece of scientific work there are people who help you along the way,” he says. “But knowing that you developed a project from start to finish largely under your own direction is gratifying. It’s a nice validation of my ability to do science.”

John Whitfield is a freelance journalist in London.

1. Wuchty, S., Jones, B. F. & Uzzi, B. *Science* **316**, 1036–1039 (2007).
2. Cummings, J. N. & Kiesler, S. *Proc. CSCW’08* (in the press).
3. Palla, G., Barabási, A.-L. & Vicsek, T. *Nature* **446**, 664–667 (2007).
4. Guimerà, R., Uzzi, B., Spira, J. & Amaral, L. A. N. *Science* **308**, 697–702 (2005).
5. Uzzi, B. *J. Phys. A* **41**, 224023 (2008).
6. Newman, M. E. J. Preprint at arXiv:0809.0522v1 (2008).
7. Börner, K., Penumarthi, S., Meiss, M. & Ke, W. *Scientometrics* **68**, 415–426 (2006).
8. Price, Derek J. de Solla *Little Science, Big Science* (Columbia Univ. Press, 1963).
9. Abt, H. A. *Scientometrics* **73**, 353–358 (2007).
10. Friedman, M. *Nature* **454**, 209–212 (2008).

What lies beneath

More creatures live in soil than any other environment on Earth. But what are they all doing there? **Amber Dance** reports on the world's widest biodiversity.

Ecosystems aren't green; they are black and brown, at least in the colour palette favoured by Diana Wall. Wall, a soil ecologist at Colorado State University in Fort Collins, spends her days digging into the world's underground ecosystems. These beiges, ochres and charcoals reflect a three-dimensional mosaic of micro-environments, each with its unique set of inhabitants.

But very little is known about these inhabitants. Understanding soil is a matter of rising urgency. A July report from the US National Research Council listed soil quality as the biggest barrier to higher crop yields for farmers in sub-Saharan Africa and south Asia. And knowing what myriad organisms live in the soil, and how they interact, is crucial to creating a healthy ecosystem.

For those scientists who are willing to crouch down and dig, the diversity of soil denizens beats any above-ground system, even that of a tropical rainforest. A handful of soil from one spot may house a very different community from soil just a metre away, because of variations in the availability of water or nutrients. For example, the ground under a decaying plant or animal is a different environment from soil lacking such enrichment. And around plant roots, specialized organisms inhabit the rhizosphere, a thin layer where roots and soil organisms interact in myriad ways. Large animals such as moles contribute, changing and aerating the underground landscape by tunnelling. Even a small clump of soil has a gradient of oxygen from its edges to the centre, and each oxygen concentration may make the perfect habitat for different kinds of creatures. "It is the most incredible zoo," says Wall.

Take that view to a larger scale, and it is possible to appreciate just how complicated

the world's soil ecosystems are. In one ongoing study, not yet published, Wall and her colleagues scooped soil cores from two sites in Alaska, one in the tundra and one in the taiga forest. Although the sites were only 400 kilometres apart, the species living there were radically different: only 18 invertebrate taxa out of an estimated 1,300 appeared in both locations. "That just blew me away," says Wall.

And that's just looking at invertebrates, not including microbes. "As far as I know, there is no environment on Earth that is more bio-

logically diverse, per unit area, than soil," says Eric Triplett, a microbiologist at the University of Florida in Gainesville. Thanks to faster, cheaper DNA sequencing, scientists are now getting a grip on what

is down there and what those organisms might be doing. That information, in turn, could help improve soil management for agriculture and forest management for conservation.

At this point, scientists don't even agree on how many creatures they are looking for. The first DNA-based estimate of soil microbial biodiversity, published in 1990, counted about 4,000 different bacterial genomes per gram of soil¹. Since then, various studies and models have pushed the number up as high as 830,000 species per gram², down to 2,000 (ref. 3), and back up again. Most

recently, Triplett and his colleagues ran 139,000 individual sequences — more than other studies have used — and came up with an estimate of 10,000 to 50,000 species per gram of soil⁴. Complicating the matter is the fact that, because so few of these species have been

described, researchers have to group similar organisms within 'operational taxonomic units', which correspond roughly but not precisely to species designations.

Valuable species

Quantifying such diversity illustrates just how much remains to be discovered, and soil scientists are teaming up to tackle the challenge. The Tropical Soil Biology and Fertility (TSBF) Institute, run by the International Center for Tropical Agriculture and headquartered in Nairobi, has united more than 300 scientists in seven countries to survey soil organisms. The project, which began in 2002, aims to identify living indicators for fertile or poor soil, and has already identified some novel organisms that could be useful to humans.

In the Veracruz rainforest, for instance, Mexican scientists have discovered *Acaulospora*, a mycorrhizal fungus that entwines with lily roots and provides water and mineral nutrients. Last spring the researchers injected *Acaulospora* into the soil of test lily plots in Benigno Mendoza, a community in Veracruz where lily bulbs are an important cash crop. As a result, this year's harvest consists of big, first-quality bulbs that match the yields gained through using inorganic fertilizer with none of the downsides of chemical treatments. Isabelle Barois, a soil ecologist at the Institute of Ecology in Xalapa and coordinator of the



EYE OF SCIENCE/SPL



A. SYRED/SPL

TSBF Mexican team, says that the fungus could eventually help replace the expensive nitrogen fertilizer and harsh agrochemicals that farmers apply to their land five or six times a year.

Global soils contain a bounty of unusual and potentially useful organisms such as *Acaulospora* — more, theoretically, than they should. Although some species are common, there are also countless taxa found in vanishingly small numbers. Many species also seem to be redundant, eating the same foods and fulfilling the same ecosystem jobs, so scientists don't quite understand why they're there at all. "There is some debate about how many species need to be present in the soil to make an ecosystem," says Wim van der Putten, an ecologist at the Netherlands Institute of Ecology in Heteren.

Heikki Setälä, an ecologist at the University of Helsinki, took on this question with experiments in which he controlled the number of animal or microbial species in artificial ecosystems. In one study⁵, he set up soil microcosms in glass jars and added fungal species: only one in some jars, and up to 43 in others. Diverse systems decomposed more organic matter — demonstrated by higher carbon dioxide production — and produced more nitrogen compounds in the soil. But that relationship held true only at the lower end of the spectrum. Six species were better than one, but 43 weren't any better than six. "It was kind of a bummer," Setälä says. "It would be nice to tell the audience that we need all the species to make the planet green and sustain it."

The explanation for the wealth of soil biodiversity, then, remains an open question. Maybe the multitudinous creatures are simply adapted for niches that humans don't yet understand. Alternatively, they could literally be waiting for a rainy day; some organisms spring into action after a storm, fire or other disturbance, and so make the ecosystem more resilient. Or perhaps those organisms are truly redundant. "We know virtually nothing

about what controls the diversity of soil communities," says soil ecologist Richard Bardgett of Lancaster University, UK.

Triplett disagrees. "I don't think it's a vast unknown," he says. "I think there are some dominant genera out there that we could learn about pretty fast." In a follow-up to his soil biodiversity survey⁶, Triplett and his colleagues found that up to around 65% of the DNA samples from soil microbes fell into known genera, which makes those genera prime candidates for further study. For example, *Chitinophaga* was prevalent in the four distinct soils tested, from Canada, Illinois, Florida and Brazil. But a PubMed search for the genus finds only ten papers on the genus (and one of those is Triplett's), highlighting the lack of work that has been done in this area.

"My dream for the future would be that you would just take a DNA sample from the soil, and then explain what species are there, and what benefits," van der Putten says. But this kind of quick DNA test is years in the future.

Setting microbes to work

For some scientists, just defining the diversity isn't enough. Triplett, for instance, wants to alter it. He envisions harnessing the nitrogen-fixing power of bacteria that form nodules on the roots of some plants, such as legumes, and convert nitrogen from the air into a form the plants can use. He thinks he could insert some of the nitrogen-fixing (*nif*) genes from the bacteria into agricultural crops — which could then collect their own nitrogen from the atmosphere and eliminate the use of artificial nitrogen fertilizer. It has already been shown that some *nif* genes can

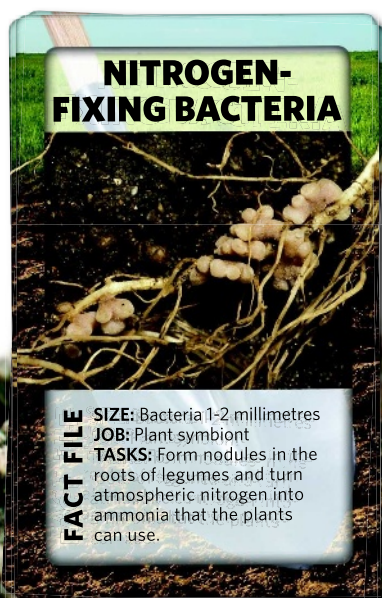
function in plants⁷. A nitrogen-fixing plant would require at least ten new genes, making the task difficult, Triplett says, but not impossible.

Policy-makers are slowly starting to pay attention to the problem of soils. In 2006 the European Union agreed that soils need protection from erosion, landslides and salinisation, but has not

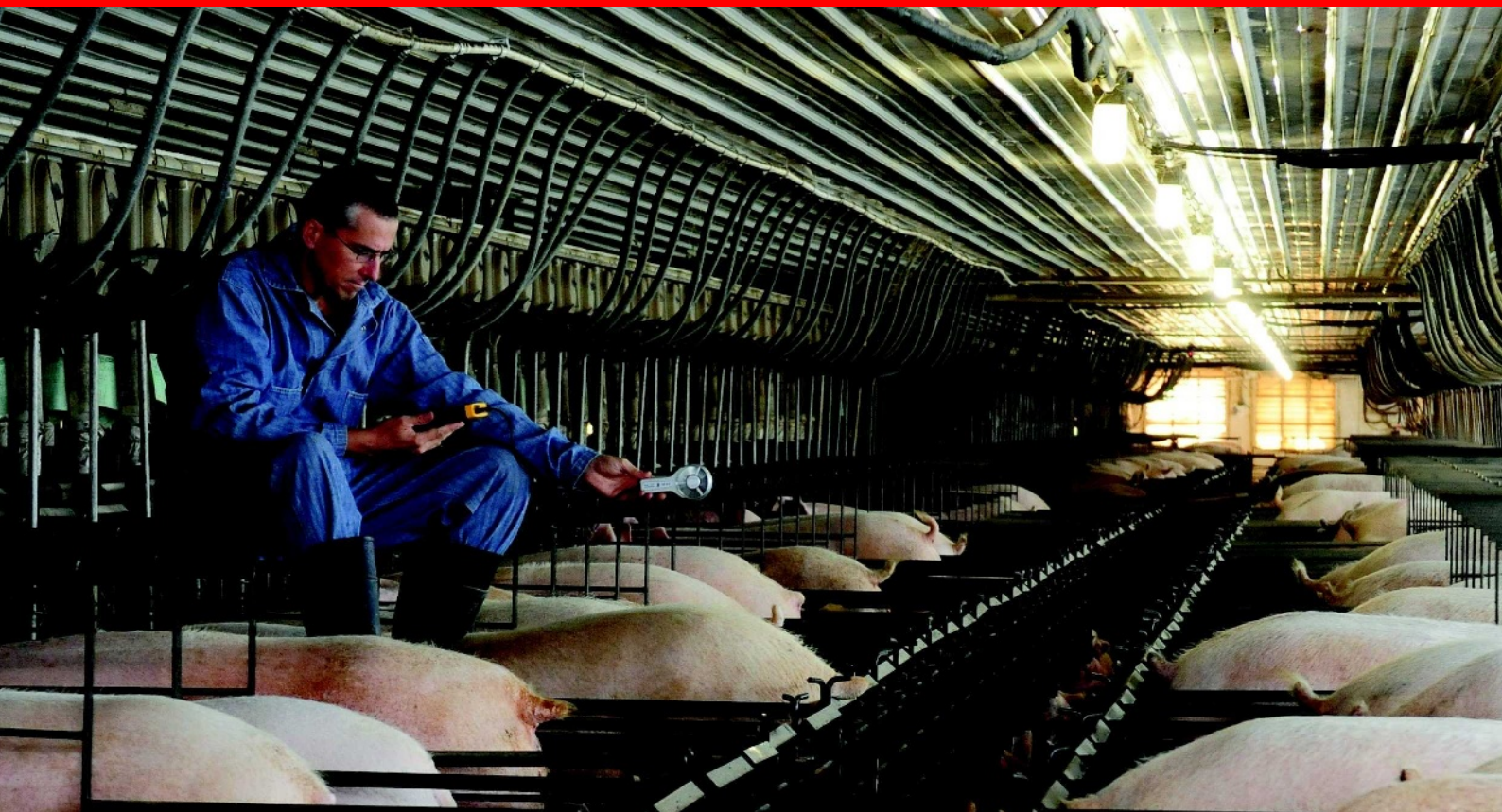
yet finalized the laws that would ensure this happens. Some countries, including France, would prefer to see individual countries regulate soil. "I'm pretty confident that the politicians will swallow the hook sooner or later," Setälä says.

Avoiding that hook comes with a price tag: one estimate valued the free services provided by the world's soil biota at US\$1.5 trillion or more each year⁸. Soils are also important as a carbon sink; soil stockpiles 1,500 gigatonnes of organic carbon, more than Earth's atmosphere and all the plants on the planet, according to the United Nations Food and Agriculture Organization. If soils remain degraded and their many denizens disappear, the world might lose access to organisms that improve crop yields, degrade toxins, or make useful by-products such as drugs — before they're even discovered. ■

Amber Dance is a freelance science writer based in the Los Angeles area, and a former News intern with Nature.



1. Torsvik, V., Goksoyr, J. & Daae, F. L. *Appl. Environ. Microbiol.* **56**, 782-787 (1990).
2. Gans, J., Wolinsky, M. & Dunbar, J. *Science* **309**, 1387-1390 (2005).
3. Schloss, P. D. & Handelsman, J. *PLoS Comput. Biol.* **2**, e92 (2006).
4. Roesch, L. F. et al. *ISME J.* **1**, 283-290 (2007).
5. Setälä, H. & McLean, M. A. *Oecologia* **139**, 98-107 (2004).
6. Fulthorpe, R. R., Roesch, L. F., Riva, A. & Triplett, E. W. *ISME J.* **2**, 901-910 (2008).
7. Cheng, Q., Day, A., Dowson-Day, M., Shen, G.-F. & Dixon, R. *Biochem. Biophys. Res. Commun.* **329**, 966-975 (2005).
8. Brussaard, L., de Ruiter, P. C. & Brown, G. G. *Agric. Ecosyst. Environ.* **121**, 233-244 (2007).



WHAT IS THAT SMELL?

Jacek Koziel is a master of odours. On a pig farm in Iowa, he shows **Erik Vance** some of the peaks and troughs of life as a human detector.

Hooking yourself up to Jacek Koziel's olfactometer is a little like having a photo album of your childhood shoved in your face. It can be offensive, sometimes painful, yet distinctly nostalgic.

Seated in Koziel's Iowa State University laboratory in Ames — with my eyes closed, and my ears covered by soundproof headphones to isolate the sense of smell — a glass cup over my nose delivers an olfactory slide show. The scents come at regular intervals. Some are strong, some less so. My job is to categorize them on a screen in front of me.

For the untrained, identifying the odours is surprisingly hard. Most of the time I get a memory rather than a name. One has a hint of lemonade and summer afternoons. Another smells like summer camp. A third reminds me of a high-school girlfriend's lip gloss.

For each one, I click a tab from the computer list in front of me with labels such as 'floral', 'cardboard' and 'rancid'. Then one hits me like an angry bull. I recoil from the machine, regain my composure, and click 'fecal'. My reaction is probably not surprising because the sample came from the gas given off by swine manure.

Koziel is an analytical chemist in the Department of Agricultural & Biosystems Engineering, where he specializes in finding and identifying trace volatile organic compounds

responsible for odour. His lab teases apart all kinds of odours, including maize (corn), wine, and the bitter fluids that insects use to defend themselves. But his "bread and butter" is livestock leavings, an incredibly complex chemical matrix. Koziel's lab is a leading authority, having catalogued almost 300 ingredients in swine manure — some of which exist only at concentrations of a few parts per million. His work is part of an international effort to understand and remove these compounds from daily life. It is a dirty job, but one that may have growing economic importance as residential areas increasingly encroach on farmland, and demand for food grows.

Koziel lives and works in a world of odour. A sign on the wall of the lab urges "good personal hygiene" while warning against wearing perfumes or scented deodorants. At the base of every whiteboard Koziel keeps scented markers to help researchers 'build awareness' of odour. And if a fume hood is accidentally left up, one quickly wonders who has broken wind.

Most people in the small, dedicated community of agricultural-odour research casually refer to odour as air pollution. But technically speaking, the United States regulates only

two of the smellier ingredients of manure — ammonia and hydrogen sulphide — for their other negative effects on soils and water sources. They are also the easiest smells to single out. "There is a tendency to go after gases that are easy to measure and historically that has been ammonia or hydrogen sulphide," says Koziel. Yet government regulations are not nearly as likely to shut down a factory farm as neighbours' complaints, which spring from a wide suite of smelly chemicals. That is because odour is the only 'air pollutant' that almost

every person living nearby has state-of-the-art detection technology for: the human nose.

It is this technology that allows Koziel to catalogue manure's most offensive components. He uses a gas chromatography mass spectrometry olfactometer (GC-MS-O). Gas chromatography separates gas samples taken from the air just above manure into different groups, based on molecular weight. Then it runs each group into two simultaneous ports. One goes to a mass spectrometer that prints out a detailed chemical analysis of the compounds. The other comes wafting out of a 'sniff port' into a researcher's waiting nose.

This blending of quantitative analysis with

"Early on it was exciting, and that's when I made a lot of faux pas at home."

— Jacek Koziel

B. ELBERT



Jacek Koziel in a gestation barn measuring air velocity and temperature.

qualitative description allows Koziel to create detailed chemical maps and find the smallest traces of compounds. Take 2-isopropyl-3-methoxy-pyrazine, a compound Koziel discovered in ladybirds that accumulated on the window sill of his old office. A mass spectrometer barely registers the compound in the cocktail of chemicals found in ladybirds. But it is unbelievably bitter stuff, which the human nose can detect in the parts per trillion. Found in wine when ladybirds get crushed with the grapes, it adds a flavour often associated with a bell pepper (*Capsicum*). Using human noses, Koziel pinpointed the offensive region of the chemical fingerprint and was able to isolate the noxious ingredient by essentially zooming in on that region (a technique called multi-dimensional GC-MS-O).

The collector

Koziel shows me a sample of the 2-isopropyl-3-methoxy-pyrazine along with a few dozen other dark bottles in a small cabinet covered with colourful warning stickers. It is housed in a sealed container, within another sealed container, within another sealed container, and is part of a frightening collection. They are the foulest odours that Koziel has come across in his work — and suffice to say, it is not a place one would want to see anything spill — just opening the wrong bottle could make the lab stink for days.

In the complex matrix of animal manure, strong but scarce compounds such as these are almost impossible to spot, yet minimizing

them may cut down on odour. Koziel's use of GC-MS-O has allowed him to create a very comprehensive view of manure that includes pungent, bitter compounds, as well as one that smells like buttered popcorn and another described as 'taco shell'.

Koziel calls himself a "smelling weirdo". Nearly anywhere he goes, he can recognize at least a few chemicals in the air. He says that distinctive flavours of body odour, red wine and mushrooms can all be found in manure. "You become more aware of the smells around you and you say 'wow, this smell smells like this chemical'. So you start making associations," he says. "Early on it was exciting, and that's when I made a lot of faux pas at home."

He once casually compared a meal his wife had made to manure. And there are the times when a family member breaks wind and he ventures to guess what they ate. Ironically, his wife says that one of the things that first attracted her to Koziel was that he smelled like the outdoors. Koziel came to the United States to climb Mount McKinley and stayed to help his climbing partner raise money for hospital bills after he lost both feet to frostbite in the attempt. It is odd to think of a mountain climber with ascents in the Andes, Indonesia and the Himalayas doing research on an Iowa pig farm, but Iowa "is a good place for odour", he says. "I have different mountains to climb, in a way."

Mountains of manure

The next day, we visit one of these mountains in the cornfields of central Iowa — a pig-birthing facility where Koziel monitors ammonia, hydrogen sulphide and dust. It is part of a nationwide detection study to quantify livestock-facility emissions. The head of that study — the National Air Emission Monitoring Study — is Al Heber. He says you cannot separate odour from air pollution. And although stink isn't a federal air pollutant, it is often the most crucial element in air-pollution debates.

"It affects proposed facilities. If that producer wants to put another farm a mile away down the road on his other property, then the odour is going to be the issue as to why he might find it difficult to locate his farm there," says Heber.

Tracking ammonia and hydrogen sulphide is a good first step, but even measuring two abundant compounds is more difficult

than it sounds. For one thing, where do you sample? Next to the farm, or next to the houses two miles away? And because odours are subject to wind and temperature, when are the most representative times to sample? And assuming you find the best time and place to sample, what is the best way to do it?

On the drive to the pig farm, not only does the smell become stronger, but it also becomes sweeter — almost like a Doppler shift of colour. That, Koziel explains, is because different odours diffuse at different rates.

Because technology cannot detect odour's reach, the most common way to measure it is dilution olfactometry — essentially, letting people smell a bit of air and asking them if it

stinks. Koziel's colleague at Iowa State University and regular co-author, Steve Hoff, is an expert in this field. He collects samples of air in specialized bags, dilutes it with odourless gas until a panel of paid

sniffers cannot detect the scent. This is the 'odour threshold'. Techniques such as this are the primary tools for quantifying odour. The problem is that often his data are essentially subjective, because one man's stink is another man's farmyard aroma.

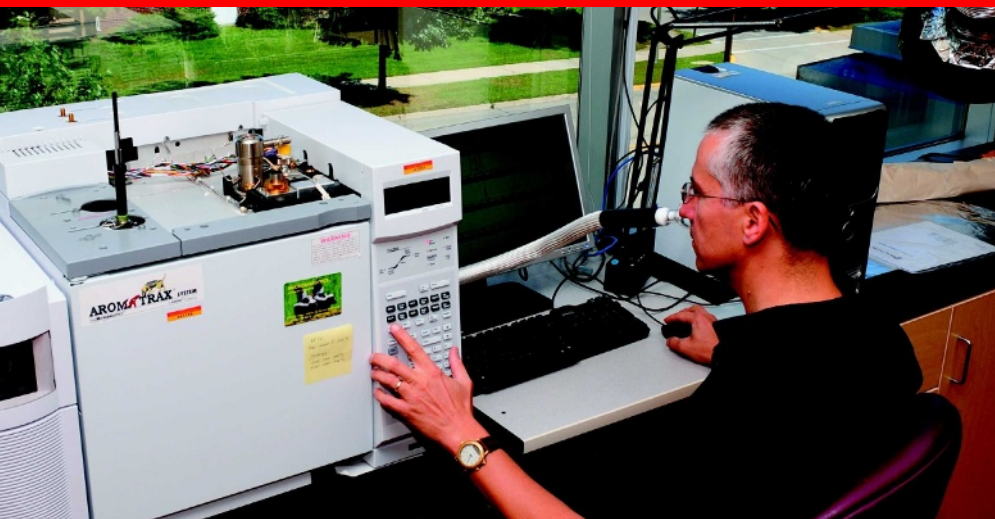
At a modern industrial factory farm, however, there are no nice ways to describe the stench. Inside the large, squat building that is Koziel's research site, 3,000 sows are either nursing piglets, having piglets or waiting to have piglets. Under the slats at my feet is one of two pits, each the size of a football field, five metres deep with manure. The sweet scent I sensed outside is now overpowered by a cacophony of others that at times make my eyes

ALAMY

"Smell affects proposed facilities. If someone wants another farm, then the odour is going to be the issue."

— Al Heber





Jacek Koziel conducts simultaneous chemical and nose 'detection' analyses of air samples.

ease, and sets about showing off a few of his hundred instruments in the building.

"This is the toughest air sampling that you can imagine. Climbing stacks or sampling smelters is okay, but this is a different ball game. You have so many variables that you have to think about," Koziel says with a rare touch of visible pride. "There is no one chimney where you just stick in your probe."

Then there's the dust. Animal dander and other particulates pump out of a pig farm at a couple of grams per cubic metre (double that for poultry). Over time this builds up, covering Koziel's detectors and every flat space in the feedlot with a millimetre or two of dust. The dust contributes to the odour's spread by carrying stowaway chemicals on the wind, but no one knows for sure to what extent.

Sticky mess

Two showers later, I get a sense of what the dust experiences. Many of the chemicals breezing about an industrial hog farm are what Koziel calls 'sticky'. That means that they are heavy molecules with a low vapour pressure that are often electrically charged. When they encounter skin, cotton or anything else with pores and crevices, they stay there. The worst will even stick to plastic or glass, making them nearly impossible to collect and measure.

These are not the easy-to-spot sulphides, but volatile fatty acids, indoles and phenols that are not a part of Heber's study. Heber says he hopes to approximate the concentration of these chemicals by measuring things such as dust and ammonia. However, in other parts of the world it's not clear that ammonia and odour are necessarily linked.

"It has an odorous component, but I don't think ammonia is really important for odour," says Nico Ogink, a farm odour scientist at the Wageningen University and Research Centre in the Netherlands where Koziel will be taking a sabbatical. "It is only important to me because of its effects on natural areas."

Ogink says that although ammonia is much more damaging to the environment, rarely does it have the ability to reach noses far from its source. This is a crucial point in the Netherlands, which has almost as many pigs as people, and where more than half of the substantial nitrogen pollution comes from livestock. In one study, 11% of the Dutch population reported being annoyed by livestock odour — more than by industrial activities and traffic.

It is difficult to say how big an effect stink has on the wallets of farmers, and industry representatives are loath to dwell on it, but several European countries such as the Netherlands, Denmark and Germany are investing heavily in odour mitigation through a series of odour-specific regulations. In Europe, the answer is

often some kind of technological filtration system. These may be bio-scrubbers, bio-filters or acid scrubbers, similar to the ones that many countries use to cut sulphur dioxide from coal-plant emissions.

The scrubbers work by exposing the air to water, which will absorb dust and suspended pollutants. European scientists say modern filters can cut odour emissions by 30–90%, depending on the conditions and whether they use several techniques together. But acid scrubbers add about €5 (US\$7) to the cost of each pig.

In the United States, where odour regulations are administered locally, mitigation may be a slurry pond out back (a practice that has led to problems in flood-prone areas) or a storage silo. Koziel's lab is working on its own mitigation techniques, most notably one that attempts to use ultraviolet light to break smells into their component parts. The process is similar to the photochemical reactions that change car exhaust into smog, only instead of creating ozone, it creates smaller, less smelly chemicals. He has also researched changes in livestock feed to try and cut the stink. Similar studies are widespread on both sides of the Atlantic Ocean. But

"Sampling in pig farms is the toughest air sampling that you can imagine."

— Jacek Koziel

Koziel is hesitant to suggest that US farmers adopt expensive mitigation approaches.

"We would be laughed at," he says. "These people work on very small margins of profit. They make money because they have a lot of animals."

When I suggest it to long-time pig farmer Chad Pierce at the Iowa pig farm, he doesn't laugh, but it is clear that's because he is trying to be polite. In good years, he can expect just \$5–10 per pig in profit. Last year, with higher feed prices, the birthing facility actually ran at a deficit. He wouldn't be opposed to smell-reducing feed, he says, but it's not a very high priority.

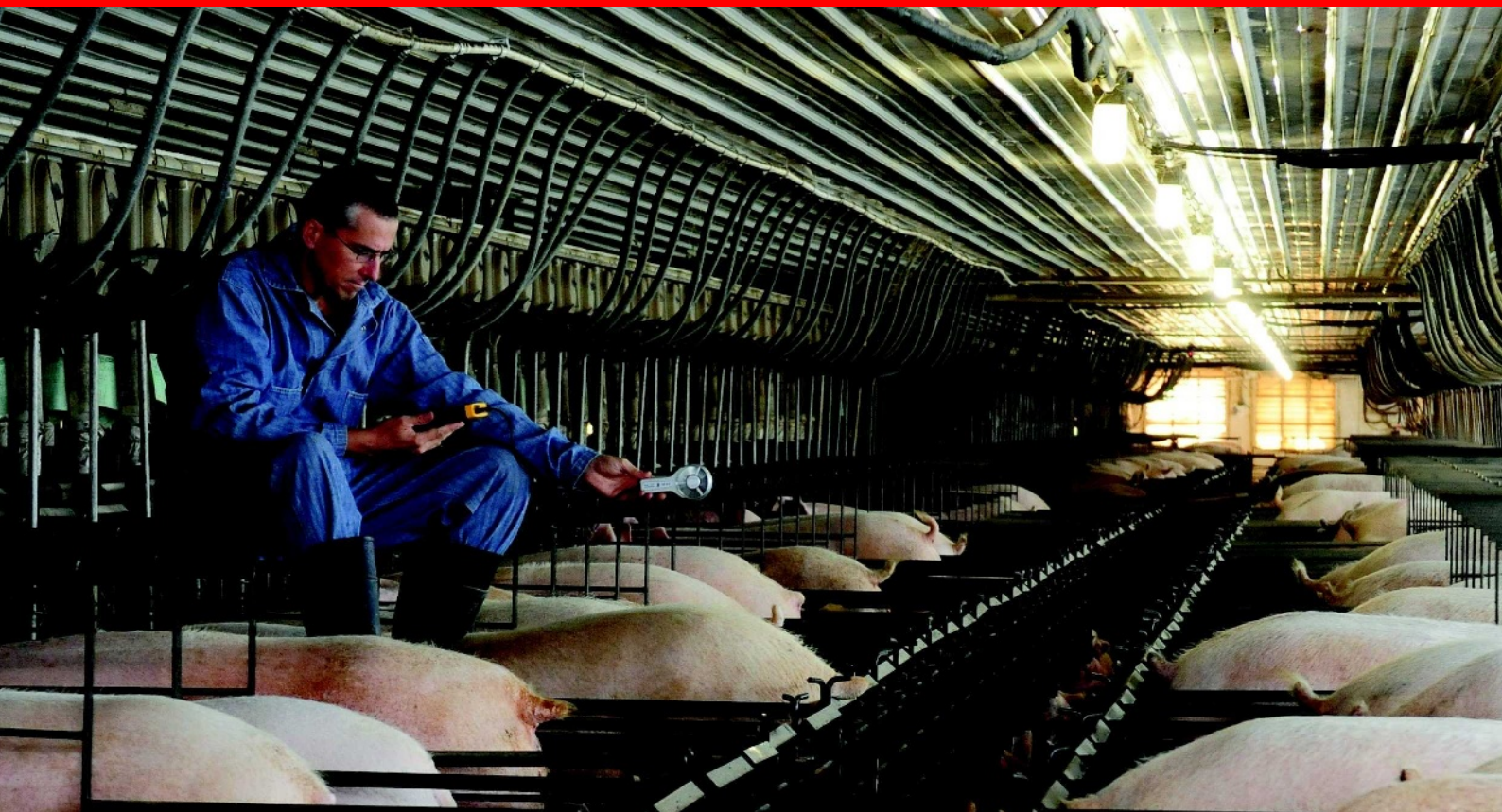
However, Pierce acknowledges that odour regulations may be on their way. And not just in the United States. One of the researchers in Koziel's lab, Lingshuang Cai, is originally from China. She says that as cities expand and citizens develop more of a taste for meats such as pork, the Chinese government is increasingly interested in odour mitigation. Although their commitment is generally ranked even lower than that of the United States, Cai says she has been offered numerous jobs in China working with manure.

But for now she is content to be in Iowa with Koziel and his mountains of smell. Stop by sometime if you are in the area. But whatever you do, don't go nosing about in the little cabinet just past the smelly fume hood — the one with the colourful warning stickers.

Erik Vance is a freelance science writer in Berkeley, California.

B. ELBERT

GETTY



WHAT IS THAT SMELL?

Jacek Koziel is a master of odours. On a pig farm in Iowa, he shows **Erik Vance** some of the peaks and troughs of life as a human detector.

Hooking yourself up to Jacek Koziel's olfactometer is a little like having a photo album of your childhood shoved in your face. It can be offensive, sometimes painful, yet distinctly nostalgic.

Seated in Koziel's Iowa State University laboratory in Ames — with my eyes closed, and my ears covered by soundproof headphones to isolate the sense of smell — a glass cup over my nose delivers an olfactory slide show. The scents come at regular intervals. Some are strong, some less so. My job is to categorize them on a screen in front of me.

For the untrained, identifying the odours is surprisingly hard. Most of the time I get a memory rather than a name. One has a hint of lemonade and summer afternoons. Another smells like summer camp. A third reminds me of a high-school girlfriend's lip gloss.

For each one, I click a tab from the computer list in front of me with labels such as 'floral', 'cardboard' and 'rancid'. Then one hits me like an angry bull. I recoil from the machine, regain my composure, and click 'fecal'. My reaction is probably not surprising because the sample came from the gas given off by swine manure.

Koziel is an analytical chemist in the Department of Agricultural & Biosystems Engineering, where he specializes in finding and identifying trace volatile organic compounds

responsible for odour. His lab teases apart all kinds of odours, including maize (corn), wine, and the bitter fluids that insects use to defend themselves. But his "bread and butter" is livestock leavings, an incredibly complex chemical matrix. Koziel's lab is a leading authority, having catalogued almost 300 ingredients in swine manure — some of which exist only at concentrations of a few parts per million. His work is part of an international effort to understand and remove these compounds from daily life. It is a dirty job, but one that may have growing economic importance as residential areas increasingly encroach on farmland, and demand for food grows.

Koziel lives and works in a world of odour. A sign on the wall of the lab urges "good personal hygiene" while warning against wearing perfumes or scented deodorants. At the base of every whiteboard Koziel keeps scented markers to help researchers 'build awareness' of odour. And if a fume hood is accidentally left up, one quickly wonders who has broken wind.

Most people in the small, dedicated community of agricultural-odour research casually refer to odour as air pollution. But technically speaking, the United States regulates only

two of the smellier ingredients of manure — ammonia and hydrogen sulphide — for their other negative effects on soils and water sources. They are also the easiest smells to single out. "There is a tendency to go after gases that are easy to measure and historically that has been ammonia or hydrogen sulphide," says Koziel. Yet government regulations are not nearly as likely to shut down a factory farm as neighbours' complaints, which spring from a wide suite of smelly chemicals. That is because odour is the only 'air pollutant' that almost

every person living nearby has state-of-the-art detection technology for: the human nose.

It is this technology that allows Koziel to catalogue manure's most offensive components. He uses a gas chromatography mass spectrometry olfactometer (GC-MS-O). Gas chromatography separates gas samples taken from the air just above manure into different groups, based on molecular weight. Then it runs each group into two simultaneous ports. One goes to a mass spectrometer that prints out a detailed chemical analysis of the compounds. The other comes wafting out of a 'sniff port' into a researcher's waiting nose.

This blending of quantitative analysis with

"Early on it was exciting, and that's when I made a lot of faux pas at home."

— Jacek Koziel

B. ELBERT



Jacek Koziel in a gestation barn measuring air velocity and temperature.

qualitative description allows Koziel to create detailed chemical maps and find the smallest traces of compounds. Take 2-isopropyl-3-methoxy-pyrazine, a compound Koziel discovered in ladybirds that accumulated on the window sill of his old office. A mass spectrometer barely registers the compound in the cocktail of chemicals found in ladybirds. But it is unbelievably bitter stuff, which the human nose can detect in the parts per trillion. Found in wine when ladybirds get crushed with the grapes, it adds a flavour often associated with a bell pepper (*Capsicum*). Using human noses, Koziel pinpointed the offensive region of the chemical fingerprint and was able to isolate the noxious ingredient by essentially zooming in on that region (a technique called multi-dimensional GC-MS-O).

The collector

Koziel shows me a sample of the 2-isopropyl-3-methoxy-pyrazine along with a few dozen other dark bottles in a small cabinet covered with colourful warning stickers. It is housed in a sealed container, within another sealed container, within another sealed container, and is part of a frightening collection. They are the foulest odours that Koziel has come across in his work — and suffice to say, it is not a place one would want to see anything spill — just opening the wrong bottle could make the lab stink for days.

In the complex matrix of animal manure, strong but scarce compounds such as these are almost impossible to spot, yet minimizing

them may cut down on odour. Koziel's use of GC-MS-O has allowed him to create a very comprehensive view of manure that includes pungent, bitter compounds, as well as one that smells like buttered popcorn and another described as 'taco shell'.

Koziel calls himself a "smelling weirdo". Nearly anywhere he goes, he can recognize at least a few chemicals in the air. He says that distinctive flavours of body odour, red wine and mushrooms can all be found in manure. "You become more aware of the smells around you and you say 'wow, this smell smells like this chemical'. So you start making associations," he says. "Early on it was exciting, and that's when I made a lot of faux pas at home."

He once casually compared a meal his wife had made to manure. And there are the times when a family member breaks wind and he ventures to guess what they ate. Ironically, his wife says that one of the things that first attracted her to Koziel was that he smelled like the outdoors. Koziel came to the United States to climb Mount McKinley and stayed to help his climbing partner raise money for hospital bills after he lost both feet to frostbite in the attempt. It is odd to think of a mountain climber with ascents in the Andes, Indonesia and the Himalayas doing research on an Iowa pig farm, but Iowa "is a good place for odour", he says. "I have different mountains to climb, in a way."

Mountains of manure

The next day, we visit one of these mountains in the cornfields of central Iowa — a pig-birthing facility where Koziel monitors ammonia, hydrogen sulphide and dust. It is part of a nationwide detection study to quantify livestock-facility emissions. The head of that study — the National Air Emission Monitoring Study — is Al Heber. He says you cannot separate odour from air pollution. And although stink isn't a federal air pollutant, it is often the most crucial element in air-pollution debates.

"It affects proposed facilities. If that producer wants to put another farm a mile away down the road on his other property, then the odour is going to be the issue as to why he might find it difficult to locate his farm there," says Heber.

Tracking ammonia and hydrogen sulphide is a good first step, but even measuring two abundant compounds is more difficult

than it sounds. For one thing, where do you sample? Next to the farm, or next to the houses two miles away? And because odours are subject to wind and temperature, when are the most representative times to sample? And assuming you find the best time and place to sample, what is the best way to do it?

On the drive to the pig farm, not only does the smell become stronger, but it also becomes sweeter — almost like a Doppler shift of colour. That, Koziel explains, is because different odours diffuse at different rates.

Because technology cannot detect odour's reach, the most common way to measure it is dilution olfactometry — essentially, letting people smell a bit of air and asking them if it

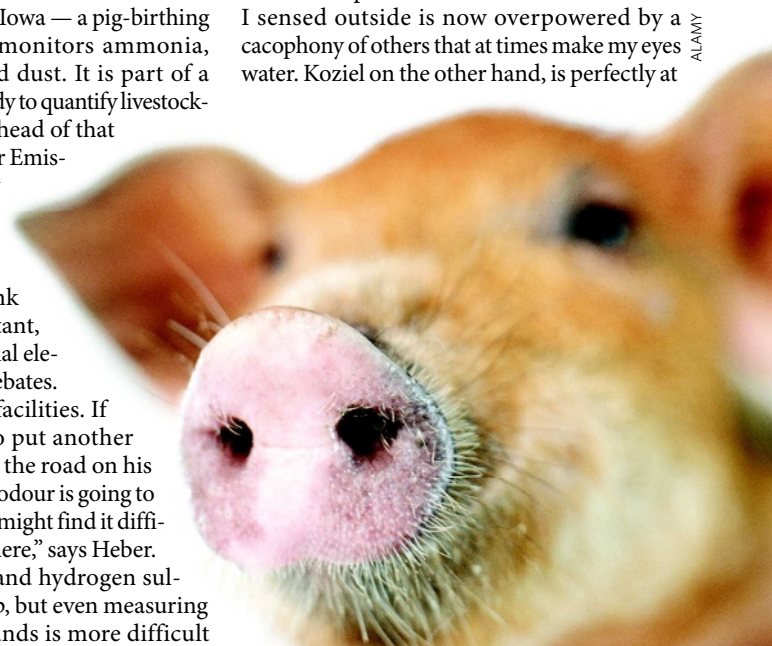
stinks. Koziel's colleague at Iowa State University and regular co-author, Steve Hoff, is an expert in this field. He collects samples of air in specialized bags, dilutes it with odourless gas until a panel of paid

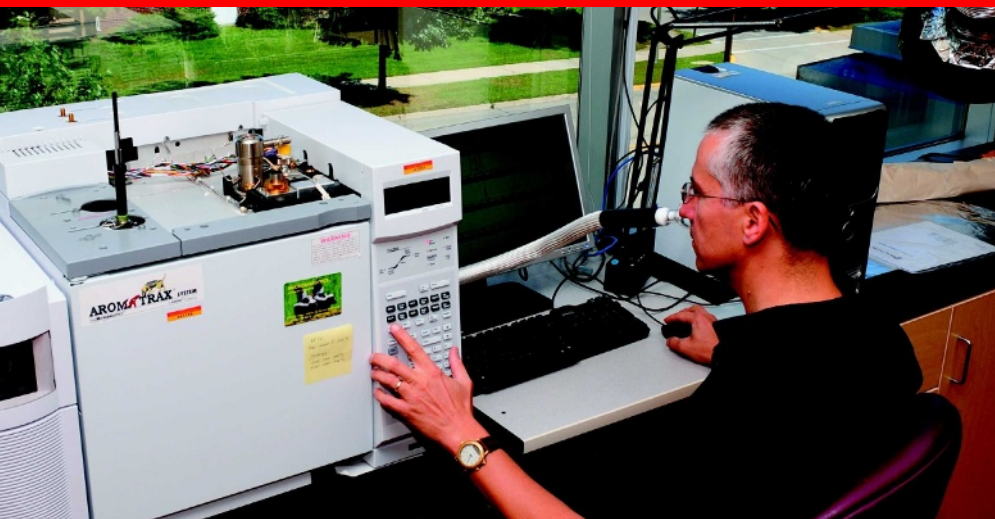
sniffers cannot detect the scent. This is the 'odour threshold'. Techniques such as this are the primary tools for quantifying odour. The problem is that often his data are essentially subjective, because one man's stink is another man's farmyard aroma.

At a modern industrial factory farm, however, there are no nice ways to describe the stench. Inside the large, squat building that is Koziel's research site, 3,000 sows are either nursing piglets, having piglets or waiting to have piglets. Under the slats at my feet is one of two pits, each the size of a football field, five metres deep with manure. The sweet scent I sensed outside is now overpowered by a cacophony of others that at times make my eyes water. Koziel on the other hand, is perfectly at

"Smell affects proposed facilities. If someone wants another farm, then the odour is going to be the issue."

— Al Heber





Jacek Koziel conducts simultaneous chemical and nose 'detection' analyses of air samples.

ease, and sets about showing off a few of his hundred instruments in the building.

"This is the toughest air sampling that you can imagine. Climbing stacks or sampling smelters is okay, but this is a different ball game. You have so many variables that you have to think about," Koziel says with a rare touch of visible pride. "There is no one chimney where you just stick in your probe."

Then there's the dust. Animal dander and other particulates pump out of a pig farm at a couple of grams per cubic metre (double that for poultry). Over time this builds up, covering Koziel's detectors and every flat space in the feedlot with a millimetre or two of dust. The dust contributes to the odour's spread by carrying stowaway chemicals on the wind, but no one knows for sure to what extent.

Sticky mess

Two showers later, I get a sense of what the dust experiences. Many of the chemicals breezing about an industrial hog farm are what Koziel calls 'sticky'. That means that they are heavy molecules with a low vapour pressure that are often electrically charged. When they encounter skin, cotton or anything else with pores and crevices, they stay there. The worst will even stick to plastic or glass, making them nearly impossible to collect and measure.

These are not the easy-to-spot sulphides, but volatile fatty acids, indoles and phenols that are not a part of Heber's study. Heber says he hopes to approximate the concentration of these chemicals by measuring things such as dust and ammonia. However, in other parts of the world it's not clear that ammonia and odour are necessarily linked.

"It has an odorous component, but I don't think ammonia is really important for odour," says Nico Ogink, a farm odour scientist at the Wageningen University and Research Centre in the Netherlands where Koziel will be taking a sabbatical. "It is only important to me because of its effects on natural areas."

Ogink says that although ammonia is much more damaging to the environment, rarely does it have the ability to reach noses far from its source. This is a crucial point in the Netherlands, which has almost as many pigs as people, and where more than half of the substantial nitrogen pollution comes from livestock. In one study, 11% of the Dutch population reported being annoyed by livestock odour — more than by industrial activities and traffic.

It is difficult to say how big an effect stink has on the wallets of farmers, and industry representatives are loath to dwell on it, but several European countries such as the Netherlands, Denmark and Germany are investing heavily in odour mitigation through a series of odour-specific regulations. In Europe, the answer is

often some kind of technological filtration system. These may be bio-scrubbers, bio-filters or acid scrubbers, similar to the ones that many countries use to cut sulphur dioxide from coal-plant emissions.

The scrubbers work by exposing the air to water, which will absorb dust and suspended pollutants. European scientists say modern filters can cut odour emissions by 30–90%, depending on the conditions and whether they use several techniques together. But acid scrubbers add about €5 (US\$7) to the cost of each pig.

In the United States, where odour regulations are administered locally, mitigation may be a slurry pond out back (a practice that has led to problems in flood-prone areas) or a storage silo. Koziel's lab is working on its own mitigation techniques, most notably one that attempts to use ultraviolet light to break smells into their component parts. The process is similar to the photochemical reactions that change car exhaust into smog, only instead of creating ozone, it creates smaller, less smelly chemicals. He has also researched changes in livestock feed to try and cut the stink. Similar studies are widespread on both sides of the Atlantic Ocean. But

"Sampling in pig farms is the toughest air sampling that you can imagine."

— Jacek Koziel

Koziel is hesitant to suggest that US farmers adopt expensive mitigation approaches.

"We would be laughed at," he says. "These people work on very small margins of profit. They make money because they have a lot of animals."

When I suggest it to long-time pig farmer Chad Pierce at the Iowa pig farm, he doesn't laugh, but it is clear that's because he is trying to be polite. In good years, he can expect just \$5–10 per pig in profit. Last year, with higher feed prices, the birthing facility actually ran at a deficit. He wouldn't be opposed to smell-reducing feed, he says, but it's not a very high priority.

However, Pierce acknowledges that odour regulations may be on their way. And not just in the United States. One of the researchers in Koziel's lab, Lingshuang Cai, is originally from China. She says that as cities expand and citizens develop more of a taste for meats such as pork, the Chinese government is increasingly interested in odour mitigation. Although their commitment is generally ranked even lower than that of the United States, Cai says she has been offered numerous jobs in China working with manure.

But for now she is content to be in Iowa with Koziel and his mountains of smell. Stop by sometime if you are in the area. But whatever you do, don't go nosing about in the little cabinet just past the smelly fume hood — the one with the colourful warning stickers.

Erik Vance is a freelance science writer in Berkeley, California.

B. ELBERT

GETTY

CORRESPONDENCE

Science lobbying in Canada needs stepping up

SIR — Your Editorial 'The other North American election' (*Nature* **455**, 263; 2008) draws attention to an apparent lack of urgency towards general science issues in the Canadian election campaign. This must surely reflect public opinion and therefore inadequate advocacy efforts by the scientific community.

By comparison with the prominence given to science-related issues in the US presidential election, their marginalization in Canadian politics is noticeable (see www.sciencecanada.blogspot.com). During the first 16 days of the election campaign's 37 days, there were only two direct pronouncements on scientific research: one by the Liberal leader on a single visit to a university and the other a mention of research and development by the prime minister during a visit to an industrial firm. That was the scale of attention to science by all main political parties: Conservatives, Liberals, New Democrats and Greens.

Science advocacy in Canada needs to be stepped up. Professional and lobbying organizations that bring science to the attention of policy-makers need to be stronger and more visible. In the United States, for example, there are numerous organizations and professional associations with dynamic and systematic links to policy-makers, Congress and the Senate; in Canada, such relations are largely conducted at a personal level.

The retired politician Preston Manning, cited in your Editorial, is not alone in calling for an independent ministry for science: many scientists have said the same. They are right to insist that science should no longer come under the ministry for industry. Its position there reflects the government's limited grasp of the importance of science policy, not

helped by the dearth of scientific and engineering training among parliamentarians.

Unlike Canada, the United States runs fellowship programmes for training young scientists in science policy and the policy-making process, through its science academies and organizations such as the American Association for the Advancement of Science.

It is time for Canadian scientists to initiate organized efforts to take science into mainstream Canadian society. This will mean strengthening existing professional organizations and establishing new and dynamic networks of science advocates. Only then can we hope that research funding will warrant more than a passing mention in political manifestos.

Mehrdad Hariri Ontario Cancer Institute, University of Toronto, 610 University Avenue, Princess Margaret Hospital, Toronto M5G 2M9, Canada
e-mail: mhariri@uhnres.utoronto.ca

Playing the system puts self-citation's impact under review

SIR — In reply to Tomáš Opatrný's Correspondence 'Playing the system to give low-impact journals more clout' (*Nature* **455**, 167; 2008), we would like to point out that the practice of journal self-citation is not new. Thomson Reuters is aware that some journals have used extensive reference to their prior content to influence their citation metrics. The contribution of so-called journal self-citation has been included in *Journal Citation Reports* since it first appeared in 1975. In recent years, these data have been made more prominent to inform our subscribers of the effects of journal self-citation.

Thomson Reuters also reviews self-citation data for journals in which an exceptionally high self-citation rate artificially

influences the impact factor and therefore belies its contribution to the scientific literature. The role of a journal's impact factor as an objective and integral measure becomes questionable at this level of self-citation.

Nine journals received no listing in *Journal Citation Reports* last year because of exceptionally high self-citation counts; their titles are listed in the Notices file on the journal's website. Journal self-citations will be reviewed each year. Once the problem of excessive self-citation resolves and we can publish an accurate impact factor, the titles will again appear in the journal. Each title continues to be indexed in other Thomson Reuters products.

The cause of the increased 2007 impact factor of *Folia Phoniatrica et Logopaedica* will be examined as part of the routine review of journal self-citations, and a decision will be made regarding continued listing of the journal in 2008's *Journal Citation Reports*.

James Testa Thomson Reuters, 3501 Market Street, Philadelphia, Pennsylvania 19104, USA
e-mail: james.testa@thomsonreuters.com

Austria: investigation likely to have serious consequences

SIR — In the Editorial 'Scandalous behaviour' (*Nature* **454**, 917–918; 2008), you unjustifiably implicate the whole Austrian nation in your comments about a scientific institution. In fact, the dismissal of the rector of the Medical University of Innsbruck had nothing to do with the alleged scientific misconduct in a urology trial that recently took place there.

The rector, Clemens Sorg, was dismissed on the grounds of unsuitability for his post after a thorough investigation by us, the university council, acting as an independent official body.

Moreover, the investigation of

the Hannes Strasser misconduct case has been top of the university council's agenda since March 2008. As you point out, the clinical urology trial has been under investigation by AGES, the Austrian agency for health and food safety — an institution comparable to the Food and Drug Administration in the United States. The agency, which issued its final report on the matter in August, routinely checks all investigator-driven clinical trials in Austria. The university council requested immediately that the rector forward the report to the state attorney — the required juridical procedure.

The ongoing investigation is likely to have serious disciplinary and legal consequences. We reject all implications that any official Austrian body is seeking to suppress this investigation.

Guenther Bonn, Christiane Druml, Gabriele Fischer, Christoph Huber, Stephan Laske, Freyja Smolle-Juettner, Richard Soyer University Council of Medical University Innsbruck, Innrain 52, 6020 Innsbruck, Austria

As someone wise or funny and probably famous once said ...

SIR — There comes a time when quotations should be laid to rest. Adam Rutherford, in his Book Review 'The future ain't what it used to be' (*Nature* **454**, 1051; 2008), cites baseball legend 'Yogi' Berra as declaring that it's tough to make predictions, especially about the future.

Maybe he did, but maybe he didn't, or at least maybe he didn't do it first. The same observation has been attributed to numerous other people, including Niels Bohr, Albert Einstein, Mark Twain, Will Rogers, George Bernard Shaw, Winston Churchill, Groucho Marx, Enrico Fermi, Freeman Dyson, Cecil B. DeMille and even Confucius (www.larry.denenberg.com/predictions.html).

Neville W. Goodman Bristol, UK
e-mail: newgoodman@mac.com

CORRESPONDENCE

Science lobbying in Canada needs stepping up

SIR — Your Editorial 'The other North American election' (*Nature* **455**, 263; 2008) draws attention to an apparent lack of urgency towards general science issues in the Canadian election campaign. This must surely reflect public opinion and therefore inadequate advocacy efforts by the scientific community.

By comparison with the prominence given to science-related issues in the US presidential election, their marginalization in Canadian politics is noticeable (see www.sciencecanada.blogspot.com). During the first 16 days of the election campaign's 37 days, there were only two direct pronouncements on scientific research: one by the Liberal leader on a single visit to a university and the other a mention of research and development by the prime minister during a visit to an industrial firm. That was the scale of attention to science by all main political parties: Conservatives, Liberals, New Democrats and Greens.

Science advocacy in Canada needs to be stepped up. Professional and lobbying organizations that bring science to the attention of policy-makers need to be stronger and more visible. In the United States, for example, there are numerous organizations and professional associations with dynamic and systematic links to policy-makers, Congress and the Senate; in Canada, such relations are largely conducted at a personal level.

The retired politician Preston Manning, cited in your Editorial, is not alone in calling for an independent ministry for science: many scientists have said the same. They are right to insist that science should no longer come under the ministry for industry. Its position there reflects the government's limited grasp of the importance of science policy, not

helped by the dearth of scientific and engineering training among parliamentarians.

Unlike Canada, the United States runs fellowship programmes for training young scientists in science policy and the policy-making process, through its science academies and organizations such as the American Association for the Advancement of Science.

It is time for Canadian scientists to initiate organized efforts to take science into mainstream Canadian society. This will mean strengthening existing professional organizations and establishing new and dynamic networks of science advocates. Only then can we hope that research funding will warrant more than a passing mention in political manifestos.

Mehrdad Hariri Ontario Cancer Institute, University of Toronto, 610 University Avenue, Princess Margaret Hospital, Toronto M5G 2M9, Canada
e-mail: mhariri@uhnres.utoronto.ca

Playing the system puts self-citation's impact under review

SIR — In reply to Tomáš Opatrný's Correspondence 'Playing the system to give low-impact journals more clout' (*Nature* **455**, 167; 2008), we would like to point out that the practice of journal self-citation is not new. Thomson Reuters is aware that some journals have used extensive reference to their prior content to influence their citation metrics. The contribution of so-called journal self-citation has been included in *Journal Citation Reports* since it first appeared in 1975. In recent years, these data have been made more prominent to inform our subscribers of the effects of journal self-citation.

Thomson Reuters also reviews self-citation data for journals in which an exceptionally high self-citation rate artificially

influences the impact factor and therefore belies its contribution to the scientific literature. The role of a journal's impact factor as an objective and integral measure becomes questionable at this level of self-citation.

Nine journals received no listing in *Journal Citation Reports* last year because of exceptionally high self-citation counts; their titles are listed in the Notices file on the journal's website. Journal self-citations will be reviewed each year. Once the problem of excessive self-citation resolves and we can publish an accurate impact factor, the titles will again appear in the journal. Each title continues to be indexed in other Thomson Reuters products.

The cause of the increased 2007 impact factor of *Folia Phoniatrica et Logopaedica* will be examined as part of the routine review of journal self-citations, and a decision will be made regarding continued listing of the journal in 2008's *Journal Citation Reports*.

James Testa Thomson Reuters, 3501 Market Street, Philadelphia, Pennsylvania 19104, USA
e-mail: james.testa@thomsonreuters.com

Austria: investigation likely to have serious consequences

SIR — In the Editorial 'Scandalous behaviour' (*Nature* **454**, 917–918; 2008), you unjustifiably implicate the whole Austrian nation in your comments about a scientific institution. In fact, the dismissal of the rector of the Medical University of Innsbruck had nothing to do with the alleged scientific misconduct in a urology trial that recently took place there.

The rector, Clemens Sorg, was dismissed on the grounds of unsuitability for his post after a thorough investigation by us, the university council, acting as an independent official body.

Moreover, the investigation of

the Hannes Strasser misconduct case has been top of the university council's agenda since March 2008. As you point out, the clinical urology trial has been under investigation by AGES, the Austrian agency for health and food safety — an institution comparable to the Food and Drug Administration in the United States. The agency, which issued its final report on the matter in August, routinely checks all investigator-driven clinical trials in Austria. The university council requested immediately that the rector forward the report to the state attorney — the required juridical procedure.

The ongoing investigation is likely to have serious disciplinary and legal consequences. We reject all implications that any official Austrian body is seeking to suppress this investigation.

Guenther Bonn, Christiane Druml, Gabriele Fischer, Christoph Huber, Stephan Laske, Freyja Smolle-Juettner, Richard Soyer University Council of Medical University Innsbruck, Innrain 52, 6020 Innsbruck, Austria

As someone wise or funny and probably famous once said ...

SIR — There comes a time when quotations should be laid to rest. Adam Rutherford, in his Book Review 'The future ain't what it used to be' (*Nature* **454**, 1051; 2008), cites baseball legend 'Yogi' Berra as declaring that it's tough to make predictions, especially about the future.

Maybe he did, but maybe he didn't, or at least maybe he didn't do it first. The same observation has been attributed to numerous other people, including Niels Bohr, Albert Einstein, Mark Twain, Will Rogers, George Bernard Shaw, Winston Churchill, Groucho Marx, Enrico Fermi, Freeman Dyson, Cecil B. DeMille and even Confucius (www.larry.denenberg.com/predictions.html).

Neville W. Goodman Bristol, UK
e-mail: newgoodman@mac.com

COMMENTARY

A network solution

With the right plan, systems biology can empower drug discovery, say **Adriano Henney** and **Giulio Superti-Furga**. Field leaders have contributed and now the authors want to hear from you.

Systems biology focuses on interactions within and between the mechanisms that combine to give rise to the function and behaviour of a biological system. To some it is the logical and inevitable next-level understanding that will propel drug discovery from empiricism to mechanism-based rational design. Countless column inches in the scientific press hail systems approaches as the latest weapon to tackle the major challenges of modern medicine. But to others, it is an ill-defined pile-up of 'omics' approaches that in terms of usefulness for drug discovery represents the culmination of all delusions. So it is not a surprise that the pharmaceutical industry remains unconvinced, fearing parallels with the genomics hype and considering the approaches currently impractical.

Can we realistically expect systems biology to have a tangible effect on human health in the near term? In June, at a workshop in Portofino, Italy,¹ a representative group of the community of academic, biotechnology and pharmaceutical scientists active in systems biology gathered with the goal of presenting a set of recommendations that, if implemented, would represent a coherent, structured route to demonstrate the value of systems biology to drug discovery in the next three to five years.

Discussions that might have resulted in defining systems biology more narrowly were deliberately avoided. We focused instead on the effect that systems approaches could have on therapeutic and diagnostic innovation, with the intention of providing recommendations and a roadmap. If implemented, this plan should provide the necessary evidence of practical utility to the industry, eventually resulting in broader, more vigorous adoption. The following recommendations stem from our review. Now we wish to invite comment and provoke wider debate through an online discussion forum (<http://network.nature.com/groups/systbiohumanhealth/forum/topics>), and ultimately to create the necessary thrust and focus to deliver the reform that is needed.

Setting data standards

Everybody at the workshop concurred on an obvious but fundamental point: systems approaches require an emphasis on quality over quantity of data. Currently researchers are flooded with unstructured, non-standard

and often semi-quantitative data that are hard to integrate into biological models. Additionally, there are multiple gaps in the data stream linking experiment to clinical outcome. For the field to advance effectively, we need to find new ways to create a reliable data pipeline that is compatible with the needs of systems biology. This pipeline can only be achieved realistically by the establishment of a consortium to set appropriate standards, quality-control metrics and processes on behalf of the community. How best to achieve such standardization procedures remains to be seen, but as soon as they are established, initial efforts should focus on closing the data gaps and on acquiring multiple data types on a limited number of standard models and samples. Lastly, it is of paramount importance to concentrate on the creation of necessary computational tools to provide a streamlined and standardized framework for handling and sharing large quantitative data sets efficiently. Interactions at Portofino highlighted a broad community willingness and motivation to apply self-discipline and adopt standards.

Modelling drug actions

Molecular pathways and networks have already been developed and used to model drug actions. For example, a computational model of the ion-channels in the heart has been used to study the role of the late sodium current in cardiac arrhythmia and to explain the previously ill-understood mode of action of the cardio-protective drug ranolazine^{2,3}. Elsewhere, maps of molecular pathways and their interconnection within larger intracellular or intercellular networks are providing a framework for the evaluation of novel therapeutic strategies. Several small molecule kinase inhibitors, such as dasatinib and bosutinib simultaneously hit several targets within a pathway. Others, such as sorafenib and sunitinib operate in different pathways, or even in different cells and processes^{4,5}. Combination therapy using drug cocktails has been successful in a number of areas including anti-HIV regimens and the treatment of rheumatoid arthritis with methotrexate and biologicals that target tumour necrosis factor. Computational approaches are being developed to model the biological

response to combination therapies⁶.

Although they are already being used extensively post-hoc in an explanatory mode to rationalize specific outcomes, these models and approaches have not yet been integrated routinely as tools in early drug discovery. Concentrating on the development and refinement of these approaches, and finding opportunities to apply them in areas, such as solid tumours and inflammation, in which other methods have proven inefficient, would help to build confidence in the ability of systems-based approaches to make predictions. We think that this step represents a good opportunity to provide early evidence of systems biology's value.

Predictive toxicology

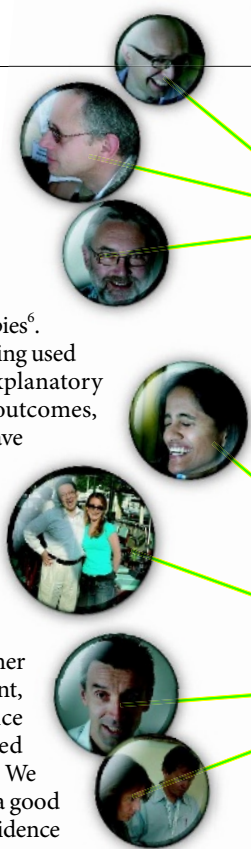
The evaluation of drug safety and toxicity is largely empirical, expensive and error prone, relying on proprietary data and often on chemistry-centric evaluations. Integrating biological data and applying quantitative modelling approaches could rapidly deliver a change in the way drug candidates are assessed. An essential first step is the assembly of as much data as possible on biological readouts linked to chemical structures (for example, of failed compounds, transcriptional signatures or metabolic profiles in cardio-, nephro- or hepato-toxicity). We hope that the pharmaceutical industry might

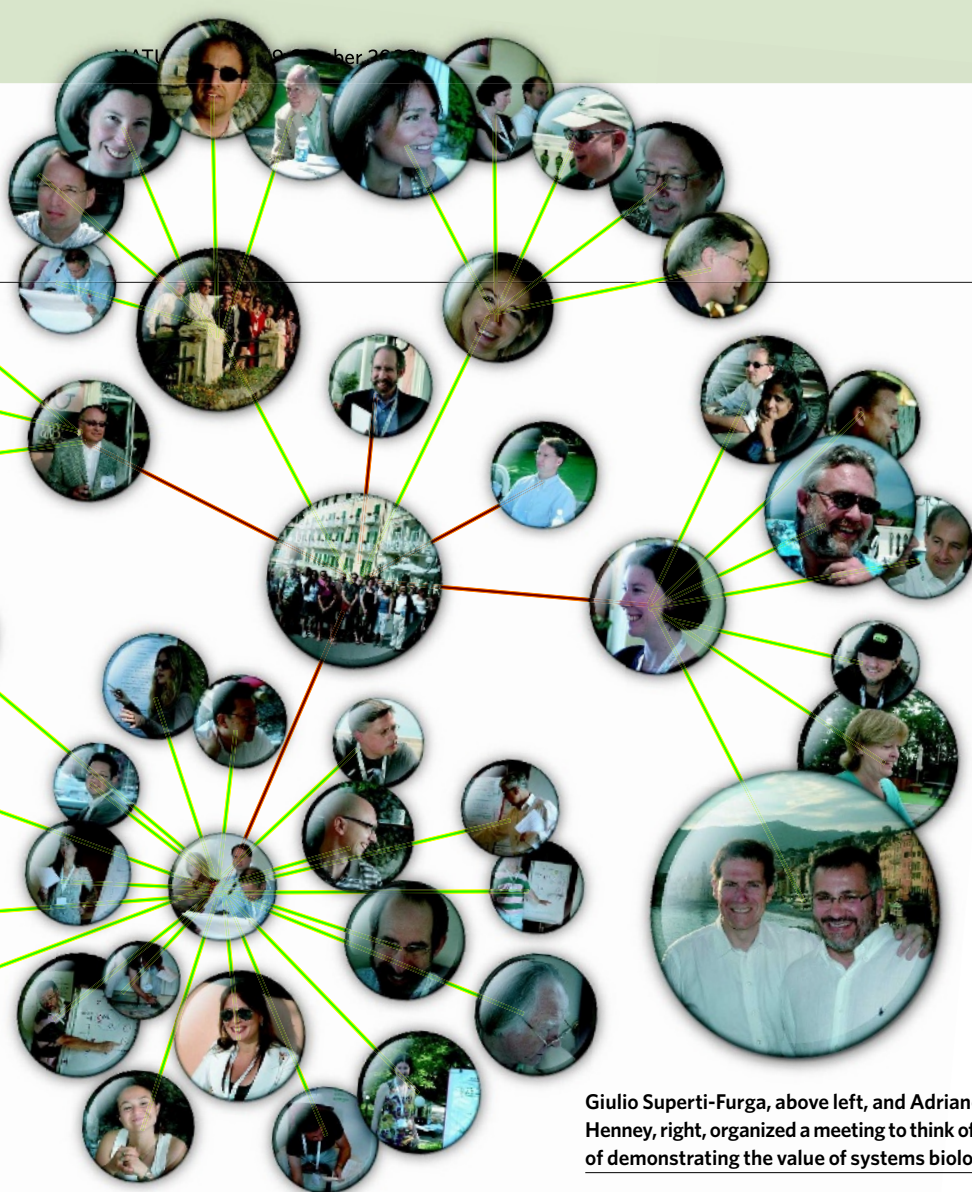
be persuaded to volunteer at least a part of its historical toxicology records to complement public-domain data in a common consortium.

A complementary step in a systems-driven approach will be the mapping of

molecular interactions relevant to toxicology. With these types of information assembled, it will be possible to create models of both human and animal, cellular and organ toxicity that can then be trained, tested and refined. This would be achieved by the recognition of molecular patterns and correlations highlighting pathways and processes elicited by chemical groups. The ultimate aspiration of developing a more comprehensive, pre-competitive predictive

"Demystifying systems biology and clarifying what the approaches can do are critical to success."





Giulio Superti-Furga, above left, and Adriano Henney, right, organized a meeting to think of ways of demonstrating the value of systems biology.

of systems-based analyses is too important to be jeopardized by a problem of image and perception, or by parochial interests and fragmented efforts. Our discussions highlighted that concerted community action is needed to ensure that the emerging success stories, and the metrics by which systems approaches will be evaluated, are explained in clear and frank terms that avoid ambiguities and hype. It is essential to understand what stakeholders really need to hear, so we can focus on the delivery of solutions rather than approaches.

As a first step towards addressing this point, this Commentary will be followed by the publication of a detailed version of the workshop discussions in the form of a white paper, incorporating comments invited from additional opinion leaders. We are also inviting broader comment and are opening the topic up for debate in an online forum. A second meeting within the next two years will review progress against the original objectives and revise or modify them as necessary in the light of this feedback, with the outcomes published using similar online tools. We hope the creation of these initiatives will lay the foundations for establishing a focused community effort in this area.

Delivering a comprehensive, integrated systems-biology toolbox for medicine and drug discovery is a 'big science' project, comparable to, but perhaps more challenging than any other major biological endeavour undertaken in the last 20 years. This challenge demands a coordinated community effort if it is ever to reach its huge potential and have a real impact. The recommendations outlined here, could be the first steps taken towards this end. ■

Adriano Henney is a director of Global Discovery Enabling Capabilities and Sciences at AstraZeneca, Alderley Park, UK.

e-mail: adriano.henney@btinternet.com

Giulio Superti-Furga directs the Center for Molecular Medicine of the Austrian Academy of Sciences in Vienna, Austria.

e-mail: gsuperti@cemm.oeaw.ac.at

1. AstraZeneca, the Biotechnology and Biological Sciences Research Council, the Centre for Molecular Medicine of the Austrian Academy of Sciences and the Wellcome Trust Workshop. *Beyond the hype: Putting systems biology to work for drug discovery*. Portofino, Italy, 9–11 June 2008.
2. Noble, D. & Noble, P. J. *Heart* **92**, 1–5 (2006).
3. Belardinelli, L., Shryock, J. C. & Fraser, H. *Heart* **92**, 6–14 (2006).
4. Rix, U. et al. *Blood* **110**, 4055–4063 (2007).
5. Bantscheff, M. et al. *Nature Biotechnol.* **25**, 1035–1044 (2007).
6. Lehár, J. et al. *Mol. Syst. Biol.* **3**, 80 (2007).
7. Mo, M. L., Jamshidi, N. & Palsson, B. Ø. *Mol. BioSyst.* **3**, 598–603 (2007).
8. Hartwell, L. H., Szankasi, P., Roberts, C. J. Murray, A. W. & Friend, S. H. *Science* **278**, 1064–1068 (1997).

To comment on this article and related documents, visit the Nature Network Group 'Systems Biology for Drug Discovery' at <http://tinyurl.com/5yj2dh>.

engine for common benefit would require full participation by industry. First, evidence of the predictive power and value of these approaches will be needed. In the meantime, even a modest improvement in the current ability to understand and quantify risk will have a significant impact on improving the safety of novel medicines. Of all areas, this could be where the economic impact could be most immediate.

Amenable therapeutic areas

Although we recognize that real impact on individual therapeutic areas is likely to take longer than the three to five years contemplated here, we nevertheless considered it important to identify those therapeutic areas best positioned to benefit from systems approaches. This was achieved through a structured, objective analysis of information collected before the workshop through interviews and questionnaires. Three main areas emerged in which a concerted community effort could deliver significant advances using the currently available data and methodology: metabolic disorders, cancer and inflammation/infectious diseases. In particular, there is a real possibility of having an effect in metabolic disorders, in which quantitative

flux human models are becoming a reality⁷. In cancer, vulnerability models can be used to look for therapies targeting cancer-type specific pathway weaknesses⁸. For inflammation and infectious diseases, modelling approaches can be important for evaluating therapeutic strategies that tackle inflammatory processes without interfering with natural host defence mechanisms. This is becoming increasingly feasible because of the understanding of bacterial, viral, fungal and parasitic systems that has emerged in recent years.

For each of the three areas, the Portofino Group sketched out an initial roadmap of activities that are needed to take us from where we are today to the delivery of a specific, stated outcome. These roadmaps (which will be made available at the online forum) are living documents that will be used to monitor and review progress towards the eventual goals. Community feedback is imperative if this is to be achieved.

Communication and outreach

Demystifying systems biology as a catch phrase and clarifying what the approaches can do are seen as critical to success, and ultimately, to influencing a cultural change in the way we approach drug discovery, moving from reductionism to integration. The real promise

BOOKS & ARTS

The many faces of carbon

An enticing new book ties together the vital roles this element has in life, the Universe and climate change, explains **John Meurig Thomas**.

The Carbon Age: How Life's Core Element Has Become Civilization's Greatest Threat

by Eric Roston

Walker & Company: 2008. 304 pp. \$25.99

Eric Roston is a journalist and science writer who covered the 9/11 attacks as a reporter for *Time* magazine. In his fascinating book *The Carbon Age*, Roston weaves together the story of the element carbon, mining his facts largely from electronic research databases, particularly Google Scholar. Providing for the layman the 'connective tissue' of a vast array of subdisciplines — encompassing anthropology, astrophysics, biotechnology, genetics, geology, mathematics, nuclear synthesis, nucleic acids, nanotechnology, palaeobotany, phylogeny and more — this US-centric monograph is a success, especially in dealing with climate change.

Roston's approach calls to mind the Christmas lectures given by Michael Faraday at the Royal Institution of Great Britain in 1848–49, and again in 1860–61, entitled "The Chemical History of a Candle". Faraday used the candle example as a door opening onto many other areas of science: "so wonderful are the varieties of outlet which it offers into the various departments of philosophy. There is not a law under which any part of this Universe is governed which does not come into play, and

is touched upon in these phenomena."

Carbon, rather than a candle, takes centre stage in Roston's attentions: its creation by nuclear synthesis in the stars, its assimilation by our planet, the generation of carbon dioxide and myriad other phenomena. Through a better understanding of these processes, he argues, we may comprehend the nature of the Universe we inhabit and find clues to overcome the problems that humans have created that bring us to the brink of global crisis.

Roston's fluent writing can be pleasing, no more so than in the chapters entitled 'CO₂ and the Tree of Life' and 'The Potential of Biological Fuels', and in his prologue. He deals lyrically with the chemical systems of the ginkgo tree, the living symbol of survival through extremes of conditions and time. However, the book contains too many US slang terms, such as "gazillions", and the text is replete with anthropomorphisms, such as "the enmity between hydrogen molecules and helium atoms". There are some oversimplifications — it is surely better to describe isotopes in the conventional manner rather than to say misleadingly that "the sum of protons and neutrons are called isotopes".

Yet *The Carbon Age* makes a compelling case concerning solutions to the problems of climate change. "Industry needs to find a way to live inside the biosphere," Roston writes.



D. SEWELL/PANOS

"Scientists are only beginning to understand biochemistry enough to find possible answers to our energy and climate crises within it." He also summarizes well the beautiful work of Frances Arnold at the California Institute of Technology, Pasadena, on the directed Darwinian evolution of proteins and carbohydrates, in which she is finding ways to convert cellulose efficiently into the biofuel, butanol.

Roston has an eye for historical perspective, as when he quotes from the evocative 1995 novel *The Rings of Saturn*, in which German émigré author W. G. Sebald visits the United Kingdom and asks how its forests, particularly at Dunwich in Suffolk, have been ravaged owing to man's preoccupation with burning. "From the earliest

Burning down the house

Global Catastrophic Risks

Edited by Nick Bostrom and Milan Ćirković

Oxford University Press: 2008. 550 pp.

\$50, £25

Eschatology, the study of how and when the world will end, has always grabbed attention, but perhaps never more so than now. Disaster movies play out our fear of our own extinction following a massive comet impact or cataclysmic volcanic super-eruption. The market for books on doom, gloom and disaster has never been so buoyant, nor so crowded — 148,000 titles on the online book shop Amazon contain the word "catastrophe".

A new addition to the genre is *Global Catastrophic Risks*, a weighty, academic tome edited by Nick Bostrom, aptly hailing from the University of Oxford's Future of Humanity

Institute, and Milan Ćirković, a professor of cosmology at Serbia's University of Novi Sad. Eschewing a tight focus, Bostrom and Ćirković pull together 23 chapters covering the complete spectrum of events that could severely damage our world and civilization. They consider the philosophical, psychological and financial aspects of global catastrophic risks, and address the usual suspects of asteroid impacts, volcanic megablasts, plague and pandemic, and nuclear war. They also touch on more exotic potential terminations at the hands of over-enthusiastic nanobot replicators, by the anthropogenic creation of exotic matter such as black holes or strange particles, or through the rise of beligerent artificial intelligence.

As an eschatologist's almanac, the book works well, providing a mine of peer-reviewed information on the great risks that threaten

our own and future generations. Yet the book is unbalanced, devoting just three chapters to natural catastrophes of global reach. Unlike technological catastrophes such as nuclear terrorism or rampant biowarfare, which may never happen, the long-term probability of occurrence of natural cataclysms such as comet impacts and volcanic super-eruptions is 100%.

More disturbing still is the downplaying of contemporary climate change, the greatest threat facing our planet, covered in a single chapter. In the editors' view, "global warming commandeers a disproportionate fraction of the attention given to global risks." Yet they also observe that "a wise person will not spend time installing a burglar alarm when the house is on fire". When your house is burning down, surely you use all the means at your disposal to put the fire out. The lack of immediate concern doesn't end there. In an otherwise insightful foreword, Martin Rees, president of the UK Royal Society,



Carbon dioxide concentrations must drop if natural disasters caused by climate change are to be prevented.

times, human civilisation has been no more than a strange luminescence growing more intense by the hours, of which no one can say when it will begin to wane and when it will fade away", Sebald ponders. "For the time being, our cities still shine through the night, and the fires still spread." These words prompt Roston to declare that "we as individuals and as a society, as nations and as species are deciding that our lifestyle is more important than its continuity".

Vivid and important passages pertaining to colourful or pioneering individuals, notably the astronomer Fred Hoyle, are well presented. In retrospect, it is clear that Hoyle should have shared the Nobel prize in physics with William A. Fowler for his contribution to the

synthesis of elemental carbon from the fusion of three helium nuclei. As well as Faraday, Roston highlights lesser known, loner scientists, such as Guy S. Callender, who realized in 1938 what was later confirmed by and credited to Charles Keeling in 1955 — that man-made CO₂ contributes to global warming. Although Roston mentions the important work on greenhouse gases by Svante Arrhenius in 1896 and John Tyndall in 1860, he does not discuss Joseph Fourier's pioneering studies in 1827.

Roston states the oft-repeated but erroneous claim that the 1985 study of carbon-60, or buckminsterfullerene, became "a founding moment for the nanotechnology movement". Not so. The contributions of Richard Feynman in 1958, Norio Taniguchi in 1974, Eric Drexler in the 1980s, and the coming of cluster science and new microscopic techniques all lay greater and prior claim. In the petrochemical industry, for example, supported nanocatalysts consisting of just a few atoms of platinum were used industrially 15 years before the carbon-60 frenzy. Even the advocacy of nanotechnology by former US president Bill Clinton takes higher priority, in influence, than buckminsterfullerene in the pervasive talk of the nanoworld.

The Carbon Age contains some minor infelicities, yet its many fine attributes win out: it is teeming with unexpected information and is a grand tour of the Universe. ■

John Meurig Thomas is professor of solid-state chemistry at the Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB2 3QZ, UK. He is author of *Principles and Practice of Heterogeneous Catalysis*. e-mail: jmt2@cam.ac.uk

deems that the main downsides of climate change lie "a century or more in the future", whereas the authors of a chapter on the threat from cosmic rays persist with the discredited idea that "current global warming may be driven by enhanced solar activity". Neither statement stands up to scientific scrutiny.

If we are to evaluate future global threats sensibly, we must distinguish between real and projected risks. We should consider separately events that are happening now, such as anthropogenic climate change and the imminent peak in the oil supply, other events that we know with certainty will occur in the longer term, notably asteroid and comet impacts and volcanic super-eruptions, and extrapolated risks, such as those associated with developments in artificial intelligence and nanotechnology, which are largely imagined.

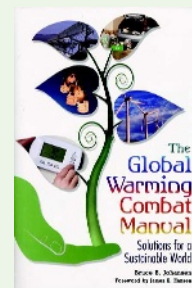
Any ranking exercise of global threats must put contemporary climate change and peak oil firmly in the first two places. Yet the latter

is not even mentioned in *Global Catastrophic Risks*. Crystal-ball gazing and horizon scanning are warranted to avoid unexpected future shocks, but these efforts should not come at the expense of ignoring the severe threats that are already staring us in the face. Closing our eyes to dangerous climate change and fast-dwindling fossil fuels will bring about a failing society that is not equipped to address any other major threats, natural or anthropogenic. To return to Bostrom and Čirković's analogy: a mushroom cloud may hang over the distant horizon and nano-goo may be oozing in our direction, but we still need to douse the flames wrought by climate change and peak oil if we are to retain for ourselves a base from which to tackle such menaces, when and if required. ■

Bill McGuire is director of the Benfield UCL Hazard Research Centre, University College London, Gower Street, London WC1E 6BT, UK. He is author of *Seven Years to Save the Planet*. e-mail: w.mcguire@ucl.ac.uk

Books on natural cataclysm, often climate-induced, are in vogue. Two books, Bruce Johansen's *The Global Warming Combat Manual* (Praeger, 2008) and Bill McGuire's *Seven Years to Save the Planet* (Weidenfeld & Nicolson, 2008), seek to make readers more environmentally aware.

McGuire explains how our homes can be made more energy efficient and our holidays greener. He recommends draconian measures to reduce China's greenhouse-gas emissions, and highlights the battle over crops for biofuels and for food production. *The Global Warming Combat Manual* discusses climate-driven changes in US environmental policy and practice. By describing recent switches — the increasing use of wind power in Texas, for example — and emerging technologies and trends, Johansen lays out a range of solutions for minimizing climate change.



CO₂ Rising by Tyler Volk (MIT Press, 2008) explains the carbon cycle in detail. Volk describes how evidence of past changes in carbon dioxide levels is obtained and interpreted, before discussing recent increases brought about by our use of fossil fuels.



Edited by Kurt Campbell, *Climatic Cataclysm* (Brookings Institution Press, 2008) looks beyond the findings of the Intergovernmental Panel on Climate Change (IPCC) to the consequences of expected, severe or even catastrophic climate change, and to what national and international organizations can do now to prevent future threats.

Dire Predictions by Michael Mann and Lee Kump (Dorling Kindersley, 2008) is an illustrated guide to the IPCC reports. Running through the scientific evidence, conclusions and key figures behind these influential reports, the book explains the latest thinking on climate science to a general reader.

Jenny Meyer

CRISIS READING



Carbon dioxide concentrations must drop if natural disasters caused by climate change are to be prevented.

times, human civilisation has been no more than a strange luminescence growing more intense by the hours, of which no one can say when it will begin to wane and when it will fade away", Sebald ponders. "For the time being, our cities still shine through the night, and the fires still spread." These words prompt Roston to declare that "we as individuals and as a society, as nations and as species are deciding that our lifestyle is more important than its continuity".

Vivid and important passages pertaining to colourful or pioneering individuals, notably the astronomer Fred Hoyle, are well presented. In retrospect, it is clear that Hoyle should have shared the Nobel prize in physics with William A. Fowler for his contribution to the

synthesis of elemental carbon from the fusion of three helium nuclei. As well as Faraday, Roston highlights lesser known, loner scientists, such as Guy S. Callender, who realized in 1938 what was later confirmed by and credited to Charles Keeling in 1955 — that man-made CO₂ contributes to global warming. Although Roston mentions the important work on greenhouse gases by Svante Arrhenius in 1896 and John Tyndall in 1860, he does not discuss Joseph Fourier's pioneering studies in 1827.

Roston states the oft-repeated but erroneous claim that the 1985 study of carbon-60, or buckminsterfullerene, became "a founding moment for the nanotechnology movement". Not so. The contributions of Richard Feynman in 1958, Norio Taniguchi in 1974, Eric Drexler in the 1980s, and the coming of cluster science and new microscopic techniques all lay greater and prior claim. In the petrochemical industry, for example, supported nanocatalysts consisting of just a few atoms of platinum were used industrially 15 years before the carbon-60 frenzy. Even the advocacy of nanotechnology by former US president Bill Clinton takes higher priority, in influence, than buckminsterfullerene in the pervasive talk of the nanoworld.

The Carbon Age contains some minor infelicities, yet its many fine attributes win out: it is teeming with unexpected information and is a grand tour of the Universe. ■

John Meurig Thomas is professor of solid-state chemistry at the Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB2 3QZ, UK. He is author of *Principles and Practice of Heterogeneous Catalysis*. e-mail: jmt2@cam.ac.uk

deems that the main downsides of climate change lie "a century or more in the future", whereas the authors of a chapter on the threat from cosmic rays persist with the discredited idea that "current global warming may be driven by enhanced solar activity". Neither statement stands up to scientific scrutiny.

If we are to evaluate future global threats sensibly, we must distinguish between real and projected risks. We should consider separately events that are happening now, such as anthropogenic climate change and the imminent peak in the oil supply, other events that we know with certainty will occur in the longer term, notably asteroid and comet impacts and volcanic super-eruptions, and extrapolated risks, such as those associated with developments in artificial intelligence and nanotechnology, which are largely imagined.

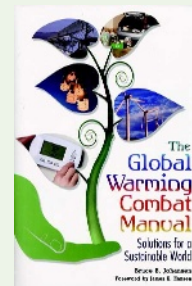
Any ranking exercise of global threats must put contemporary climate change and peak oil firmly in the first two places. Yet the latter

is not even mentioned in *Global Catastrophic Risks*. Crystal-ball gazing and horizon scanning are warranted to avoid unexpected future shocks, but these efforts should not come at the expense of ignoring the severe threats that are already staring us in the face. Closing our eyes to dangerous climate change and fast-dwindling fossil fuels will bring about a failing society that is not equipped to address any other major threats, natural or anthropogenic. To return to Bostrom and Čirković's analogy: a mushroom cloud may hang over the distant horizon and nano-goo may be oozing in our direction, but we still need to douse the flames wrought by climate change and peak oil if we are to retain for ourselves a base from which to tackle such menaces, when and if required. ■

Bill McGuire is director of the Benfield UCL Hazard Research Centre, University College London, Gower Street, London WC1E 6BT, UK. He is author of *Seven Years to Save the Planet*. e-mail: w.mcguire@ucl.ac.uk

Books on natural cataclysm, often climate-induced, are in vogue. Two books, Bruce Johansen's *The Global Warming Combat Manual* (Praeger, 2008) and Bill McGuire's *Seven Years to Save the Planet* (Weidenfeld & Nicolson, 2008), seek to make readers more environmentally aware.

McGuire explains how our homes can be made more energy efficient and our holidays greener. He recommends draconian measures to reduce China's greenhouse-gas emissions, and highlights the battle over crops for biofuels and for food production. *The Global Warming Combat Manual* discusses climate-driven changes in US environmental policy and practice. By describing recent switches — the increasing use of wind power in Texas, for example — and emerging technologies and trends, Johansen lays out a range of solutions for minimizing climate change.



CO₂ Rising by Tyler Volk (MIT Press, 2008) explains the carbon cycle in detail. Volk describes how evidence of past changes in carbon dioxide levels is obtained and interpreted, before discussing recent increases brought about by our use of fossil fuels.



Edited by Kurt Campbell, *Climatic Cataclysm* (Brookings Institution Press, 2008) looks beyond the findings of the Intergovernmental Panel on Climate Change (IPCC) to the consequences of expected, severe or even catastrophic climate change, and to what national and international organizations can do now to prevent future threats.

Dire Predictions by Michael Mann and Lee Kump (Dorling Kindersley, 2008) is an illustrated guide to the IPCC reports. Running through the scientific evidence, conclusions and key figures behind these influential reports, the book explains the latest thinking on climate science to a general reader.

Jenny Meyer

Painting by night

Van Gogh and the Colors of the Night

The Museum of Modern Art, New York
Until 5 January 2009 and then at the
Van Gogh Museum, Amsterdam, from 13
February to 7 June 2009.

Van Gogh and the Colors of the Night, an exquisite exhibition now on show at the Museum of Modern Art in New York, explores this artist's fascination with portraying the night. Unlike earlier artists who painted night scenes by day from memory, Vincent van Gogh painted his nocturnal scenes on the spot, using gaslight and colour in innovative ways to depict sunset and starlight in luminous yellow tones. As the exhibition's curator Joachim Pissarro explains, van Gogh was the first artist "to blend together gaslight — artificial, urban light — with starry light in a painting". Both lights blaze with "the same kind of buzzing, bursting of energy, a kind of weird kinship".

That "weird kinship" is captured perfectly in van Gogh's 1888 work *The Starry Night Over the Rhône* (pictured), his first attempt to paint the stars. The Great Bear constellation burns green in the sky, while across the river, the distant orange gas lamps of the French town of Arles are reflected in the water like the stars' earthly companions. Gas lamps gleam with the same fierce energy in *The Night Café*, a sparsely occupied saloon that van Gogh painted during three sleepless nights in 1888. The glare of the four gas lamps hanging from the ceiling combines

with the painting's high-pitched palette of greens and reds — six or seven shades from "blood-red to delicate pink" — to evoke, as the artist wrote, "the terrible human passions". The melancholy of this scene contrasts sharply with the joyful hubbub of *The Dance Hall in Arles* (1888), exhibited here beside *The Night Café* for the first time; the golden orbs of the gas lamps bathe the dancers in a warm and vibrant light.

These bright, spirited scenes seem eons away from van Gogh's earlier, muddy-dark depictions of country cottages and peasants. In his first significant interior night scene, *The Potato Eaters*, painted in 1885, a single oil lamp casts a pallid glow over the rough faces of the farmers as they share their meagre meal. In *The Cottage*, painted the same year, a narrow gash

of sunset sky and the smudge of an oily flame in a window are all that animate the green-black evening shadows. The light of oil lamps may also be glimpsed inside the houses of *The Starry Night*, the exhibition's most magnificent work. Van Gogh painted it in June 1889 — a year before his suicide — while confined in an asylum at Saint-Rémy in the south of France. In this delirious vista, clouds churn, the crescent Moon shines like a roiling Sun, and tall cypress trees tower in the foreground as symbols of death and the afterlife. Beneath them, the houses of the village stand serenely, their windows squares of comforting yellow light — beacons of the life inside. ■

Josie Glausiuz is a journalist based in New York.
e-mail: josiegz@earthlink.net



H. LEWANDOWSKI/REUNION DES MUSÉES NATIONAUX/ART RESOURCE

Science at the movies

Imagine Science Film Festival

New York City, New York
16–25 October 2008

CinemaScience

Village CinemaScience, Bordeaux, France
16–26 October 2008

When scientists appear on the big screen, if at all, they tend to be going mad or else paying for their hubris — think *Dr. Strangelove*, *Jurassic Park* and *A Beautiful Mind*. This month, two new film festivals — the Imagine Science Film Festival in New York (www.imaginesciencefilms.com) and CinemaScience in Bordeaux, France (www.cnrs.fr/cinemascience) — aim to correct this impression. Privileging fiction over documentary, they show how to tell stories grounded in real research.

The Imagine festival, sponsored by *Nature*, began as a series of screenings at New York's

Rockefeller University by biologist Alexis Gambis. Illness is the villain in many of the chosen short films, from Jen Peel's medical thriller *Muerto Canyon*, about a deadly virus in New Mexico, to Toddy Burton's *The Aviatrix*, about the superhero alter-ego of a woman struggling with cancer. Some of the best use humour. *California King*, directed by Eli Kaufman, is the tale of a mattress salesman who falls for an insomniac, and it is leavened with ironically placed lessons in Newtonian mechanics. Like the wordless opening of Disney-Pixar's *WALL-E*, the post-apocalyptic Pygmalion story *Lone*, from Andrew Nowrojee, has some of the pathetic charm of Buster Keaton.

A pair of pulse-quickeners features in Spanish round out the programme: *La Habitación de Fermat* (*Fermat's Room*; 2007), a thriller about a group of mathematicians forced to solve word problems or die, and Alex Rivera's

Sleep Dealer (2008), in which virtual labour and water politics make for a Mexican *Star Wars* with a Marxist twist. The festival also screens Paul Devlin's stranger-than-fiction documentary *BLAST!* (2006), about astrophysicists travelling to Antarctica to launch a telescope on a high-altitude balloon.

The CinemaScience festival in Bordeaux is sponsored by the CNRS, France's basic-research agency. The festival examines Hollywood's reliance on scientific innovation as a source of disaster, with retrospectives of classics from Fritz Lang's *Metropolis* (1927) to James Cameron's *Terminator II* (1991). As a mild corrective, Exodus Film Group's new animated feature *Igor*, about a hunchbacked lab assistant who hopes to win the Evil Science Fair, promises to poke gentle fun at common misperceptions. Other films engage more seriously with the history of science. The biopic *Korolev* (2007), directed by Yuri Kara, follows the career of the Russian rocket scientist who survived Stalin's labour camps to launch Sputnik

Painting by night

Van Gogh and the Colors of the Night

The Museum of Modern Art, New York
Until 5 January 2009 and then at the
Van Gogh Museum, Amsterdam, from 13
February to 7 June 2009.

Van Gogh and the Colors of the Night, an exquisite exhibition now on show at the Museum of Modern Art in New York, explores this artist's fascination with portraying the night. Unlike earlier artists who painted night scenes by day from memory, Vincent van Gogh painted his nocturnal scenes on the spot, using gaslight and colour in innovative ways to depict sunset and starlight in luminous yellow tones. As the exhibition's curator Joachim Pissarro explains, van Gogh was the first artist "to blend together gaslight — artificial, urban light — with starry light in a painting". Both lights blaze with "the same kind of buzzing, bursting of energy, a kind of weird kinship".

That "weird kinship" is captured perfectly in van Gogh's 1888 work *The Starry Night Over the Rhône* (pictured), his first attempt to paint the stars. The Great Bear constellation burns green in the sky, while across the river, the distant orange gas lamps of the French town of Arles are reflected in the water like the stars' earthly companions. Gas lamps gleam with the same fierce energy in *The Night Café*, a sparsely occupied saloon that van Gogh painted during three sleepless nights in 1888. The glare of the four gas lamps hanging from the ceiling combines

with the painting's high-pitched palette of greens and reds — six or seven shades from "blood-red to delicate pink" — to evoke, as the artist wrote, "the terrible human passions". The melancholy of this scene contrasts sharply with the joyful hubbub of *The Dance Hall in Arles* (1888), exhibited here beside *The Night Café* for the first time; the golden orbs of the gas lamps bathe the dancers in a warm and vibrant light.

These bright, spirited scenes seem eons away from van Gogh's earlier, muddy-dark depictions of country cottages and peasants. In his first significant interior night scene, *The Potato Eaters*, painted in 1885, a single oil lamp casts a pallid glow over the rough faces of the farmers as they share their meagre meal. In *The Cottage*, painted the same year, a narrow gash

of sunset sky and the smudge of an oily flame in a window are all that animate the green-black evening shadows. The light of oil lamps may also be glimpsed inside the houses of *The Starry Night*, the exhibition's most magnificent work. Van Gogh painted it in June 1889 — a year before his suicide — while confined in an asylum at Saint-Rémy in the south of France. In this delirious vista, clouds churn, the crescent Moon shines like a roiling Sun, and tall cypress trees tower in the foreground as symbols of death and the afterlife. Beneath them, the houses of the village stand serenely, their windows squares of comforting yellow light — beacons of the life inside. ■

Josie Glausiuz is a journalist based in New York.
e-mail: josiegz@earthlink.net



H. LEWANDOWSKI/REUNION DES MUSÉES NATIONAUX/ART RESOURCE

Science at the movies

Imagine Science Film Festival

New York City, New York
16–25 October 2008

CinemaScience

Village CinemaScience, Bordeaux, France
16–26 October 2008

When scientists appear on the big screen, if at all, they tend to be going mad or else paying for their hubris — think *Dr. Strangelove*, *Jurassic Park* and *A Beautiful Mind*. This month, two new film festivals — the Imagine Science Film Festival in New York (www.imaginesciencefilms.com) and CinemaScience in Bordeaux, France (www.cnrs.fr/cinemascience) — aim to correct this impression. Privileging fiction over documentary, they show how to tell stories grounded in real research.

The Imagine festival, sponsored by *Nature*, began as a series of screenings at New York's

Rockefeller University by biologist Alexis Gambis. Illness is the villain in many of the chosen short films, from Jen Peel's medical thriller *Muerto Canyon*, about a deadly virus in New Mexico, to Toddy Burton's *The Aviatrix*, about the superhero alter-ego of a woman struggling with cancer. Some of the best use humour. *California King*, directed by Eli Kaufman, is the tale of a mattress salesman who falls for an insomniac, and it is leavened with ironically placed lessons in Newtonian mechanics. Like the wordless opening of Disney-Pixar's *WALL-E*, the post-apocalyptic Pygmalion story *Lone*, from Andrew Nowrojee, has some of the pathetic charm of Buster Keaton.

A pair of pulse-quickeners features in Spanish round out the programme: *La Habitación de Fermat* (*Fermat's Room*; 2007), a thriller about a group of mathematicians forced to solve word problems or die, and Alex Rivera's

Sleep Dealer (2008), in which virtual labour and water politics make for a Mexican *Star Wars* with a Marxist twist. The festival also screens Paul Devlin's stranger-than-fiction documentary *BLAST!* (2006), about astrophysicists travelling to Antarctica to launch a telescope on a high-altitude balloon.

The CinemaScience festival in Bordeaux is sponsored by the CNRS, France's basic-research agency. The festival examines Hollywood's reliance on scientific innovation as a source of disaster, with retrospectives of classics from Fritz Lang's *Metropolis* (1927) to James Cameron's *Terminator II* (1991). As a mild corrective, Exodus Film Group's new animated feature *Igor*, about a hunchbacked lab assistant who hopes to win the Evil Science Fair, promises to poke gentle fun at common misperceptions. Other films engage more seriously with the history of science. The biopic *Korolev* (2007), directed by Yuri Kara, follows the career of the Russian rocket scientist who survived Stalin's labour camps to launch Sputnik

into orbit. Andrzej Wajda's acclaimed film *Katyn* (2007), about the Soviet massacre of Polish troops in 1940, is informed by a forensic investigation of their mass grave. But not all is dark and Slavic. The French comedy *La Très Très Grande Entreprise* (directed by Pierre Jolivet, on general release next month),

about workers who sue an agrochemical company for polluting their pond, is *Erin Brockovich* played for laughs. And Chilean Ricardo Larraín's *Chile Puede* (2008) tells the story of an unfortunate cosmonaut stranded in space by his own countrymen.

The festivals show that there are many ways

to get research right at the movies. "When you make a film, you want the science to be wrapped around a story," said Gambis of the Imagine festival. "I don't think you have to distort science to make it exciting."

Jascha Hoffman is a writer based in New York.
e-mail: jascha@jaschahoffman.com

Neutrinos: wonderful or crass?

John Updike's 1960 poem vents frustration at the detached and elusive quantum particles that have no consideration for our earthly perceptions, says **Martin Kemp**.

The torrent of words and images triggered by the huge publicity campaign accompanying the launch of the Large Hadron Collider at CERN, the particle-physics laboratory near Geneva, Switzerland, has served to stress the gulf between cutting-edge physics and public understanding. Terms such as 'quantum leap' and 'Big Bang' have entered our popular vocabulary. Many people are able to recognize the appearance of particle collisions from the familiar etched repertoire of lines, arcs, spirals and fusillades of points. But whether most of us can really visualize these phenomena in terms that relate to our experience is another matter.

We are evolutionarily equipped to master visible and touchable space to an extraordinary degree. But we have always experienced difficulty with the immensely big and minutely small, and the infinite lies definitively beyond the reach of our standard perceptions. In the Middle Ages, this dilemma was resolved by distinguishing between the indefinitely large and unknowable infinity, which was identified with God. The mathematician Georg Cantor still used this basic distinction in the late 1880s when he protested that his set theory was not in conflict with theology.

Finite reason and infinite revelation could each occupy their own territories in the traditional Christian solution. Something of this separation is signalled in John Updike's neat aphorism "Gods do not answer letters" in a 1960 article in *The New Yorker* — the god in this case was the baseball star Ted Williams.

The worlds disclosed by physics and astronomy have progressively taken us into realms where the separation of indefiniteness and infinity is no longer sustainable. They have also taken us into intellectual oceans where non-specialists swim in a sea of cosmic bewilderment. This sense is perfectly captured in Updike's poem *Cosmic Gall*, published in *The New Yorker* on 17 December 1960, and now often billed as the physicist's favourite poem.

Updike, so prolific and varied as to resist ready encapsulation, is famed for his refined,

lyrical and forensic treatment of small communities, centred on the fictional town of Alton, Pennsylvania. His 'Rabbit' series of novels and his 1984 book *The Witches of Eastwick*, later made into a film, have reached a public beyond the specialist world of high literature.

Cosmic Gall is both well informed and infused by an underlying irritation about the

Cosmic Gall

by John Updike

Neutrinos they are very small.
They have no charge and have no mass
And do not interact at all.
The earth is just a silly ball
To them, through which they simply pass,
Like dustmaids down a drafty hall
Or photons through a sheet of glass.
They snub the most exquisite gas,
Ignore the most substantial wall,
Cold-shoulder steel and sounding brass,
Insult the stallion in his stall,
And, scorning barriers of class,
Infiltrate you and me! Like tall
And painless guillotines, they fall
Down through our heads into the grass.
At night, they enter at Nepal
And pierce the lover and his lass
From underneath the bed — you call
It wonderful; I call it crass.

counter-experiential nature of modern physics. It concerns the neutrino, the particle proposed by Wolfgang Pauli in 1930 and subsequently named by Enrico Fermi. Its existence was ingeniously adduced by Pauli to save the laws of conservation of energy and momentum in the face of energy loss in the products of radioactive β -decay. The particle's presence had not been observed, but Pauli speculated that its lack of mass and disinclination to interact would result in its undetectability.

We now know a good deal more about the ubiquitous but furtive neutrino particles, not least courtesy of three sets of Nobel prize winners: Leon Lederman, Melvin Schwartz

and Jack Steinberger in 1988; Martin Perl and Frederick Reines in 1995; and Raymond Davies and Masatoshi Koshihara in 2002. We know that neutrinos have a very small amount of mass, that they acquire three 'flavours' that are interchangeable through oscillations, and that about 10^{14} neutrinos from the Sun and 10^3 neutrinos in cosmic rays pass through our bodies each second.

Their elusiveness is part of their fascination. In the anthropometric vocabulary that lurks in the literature of nuclear physics, neutrinos feature as 'ghosts', 'poltergeists' and 'phantoms' — to say nothing of the 'personalities' they are accorded in physicists' conversations.

Updike opens with a nursery rhyme or limerick sentence, and then bounces through the poem with a series of rapid repetitions — mass, pass, glass, gas, class, lass, crass, all, ball, hall, wall, stall, tall, fall, call and even Nepal. The apparent randomness of his examples highlights the lack of discrimination and extreme disinterestedness of the wandering neutrinos.

We sense a perverse degree of relish for the neutrinos' detachment from our daily realities, our Earth-bound perceptions and engrossing passions — and those of the stallions in their stalls. "Dustmaids down a drafty hall" is a suggestively rich image, relating to no known job description but clear in meaning by analogy to dustmen.

But should physicists really like the poem? It seems to reverse their enchanted perspective. Its 19 lines are suffused by Updike's frustration at the impersonal way that neutrinos take no notice of us, living their own non-interactive lives, if such a life is a life at all. At the end, he contrasts the scientists' wonder with his own sense that such neutrality is crass.

Perhaps a physicist who has pinned the poem on the laboratory wall can take comfort from the potential ambiguity of the "It" in the last line. Maybe the "lover and his lass" are crass. But given Updike's human values, I think not.

Martin Kemp is research professor in the history of art at the University of Oxford, Oxford OX11PT, UK.

Hidden treasures: Turin's anatomy museum

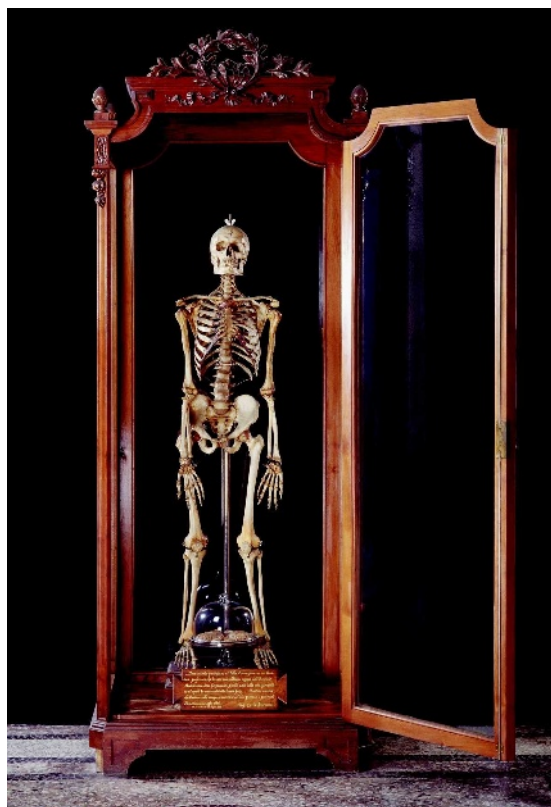
Some controversial nineteenth-century theories about brain shape and human nature are revealed by an extensive collection of neuroscience memorabilia, reports **Alison Abbott**.

Cesare Lombroso's contemporaries considered him to be either a genius or a madman, or both. A professor at the University of Turin, Italy, from 1876 until his death in 1909, Lombroso was a pioneer of criminal psychology. But his unshakeable theory that criminals were born rather than made, and could be recognized by their atavistic physical characteristics, went down badly in some quarters.

His other unshakeable theory held, ironically, that genius and madness were two sides of the same degenerate coin. In 1897, at the height of his fame, Lombroso travelled to Leo Tolstoy's village in Russia to gather living proof of the theory — but the undisputed genius disappointed him by lacking the physical characteristics that Lombroso associated with madness. In turn, Tolstoy dismissed his visitor as “ingenuous and limited”, and later described Lombroso's theories as a “misery of thought, of concept and of sensibility” (see *Nature* 409, 983; 2001). The great French novelist Émile Zola levelled that Lombroso gathered proof selectively: “like all men with pre-conceived theses.”

The irrepressible Lombroso also had plenty of opponents back home in Turin — most notably the neuroanatomist Carlo Giacomo, head of the University of Turin's anatomy museum. In the 1880s, Giacomo had developed a ‘dry’ method for preserving brains based on mummification, which he put to lavish use. At least 950 of the resulting specimens are displayed in the Museum of Human Anatomy of the University of Turin, which reopened last year after renovation, having been closed for more than a century. Giacomo was a thorough, systematic scientist interested in individual variability in the gross anatomy of the brain. His analysis of the crevices, or sulci, of human brains suggested that there is sufficient variability among normal people to negate Lombroso's theory that the size and shape of a brain dictate character. Typically, Lombroso ignored the data.

The new anatomy museum is designed to recreate the atmosphere of the original, established at Turin's ancient university in 1739 and first opened to the public by Luigi Rolando in 1830. The design is church-like, in deference to the prevailing spirit of positivism in



Closet genius: Giacomo's skeleton and unusually shaped brain.

nineteenth-century Turin, which considered science to be the new religion. The overcrowded glass cabinets are laid out in bays akin to side chapels, with paintings of famous anatomists — including Andreas Vesalius and Marcello Malpighi — hanging like saints in their lunettes. Four marble fonts previously contained floating anatomical preparations instead of holy water. There is even an 1897 stained-glass window depicting brain slices that Giacomo had prepared for microscopy.

During the revolutionary years of the early 1800s, when Napoleon occupied Turin, Rolando retreated to Sardinia, stopping en route to study anatomical wax modelling and scientific drawing in Florence. During his seven isolated years in Sardinia, he developed modern methods for studying the brain, integrating new data from anatomy, physiology and embryology. He is perhaps best known for recognizing the fixed patterns of sulci in the brain that fuelled the theories of Lombroso and Giacomo. The brain's deep central

sulcus, which separates the primary motor cortex from the sensory cortex, is named the ‘Rolando sulcus’ after him. When Napoleon lost the war in 1814, Rolando returned to Turin, where he energetically applied his new learning both to his research and to the museum.

The museum has many valuable objects, but its neuroscience collection is exceptional. Aside from the staggering piles of preserved brains provided by Giacomo, there are models of the brain created throughout the nineteenth century, which show how knowledge of brain anatomy evolved extensively during this period. The collection also includes a beautiful model from 1883 which, although rather abstract, indicates for the first time that the brain is composed of nerve pathways. There is also a collection of skulls and death masks — of the great and good, and of criminals. An exquisite series of wax human embryos shows the development of the nervous system.

Giacomo, like Rolando, lived through turbulent times. He participated in the struggles to liberate parts of northern Italy from the Austrians, one of the conflicts that eventually led to the unification of Italy in 1870. During the bloody process, Turin lost its status as capital city and needed to find a new strategy for its future. It chose science, and massively expanded its university. With the burgeoning modern industries, Turin also became Italy's film capital, and museum visitors can watch the first movies of living cells made in the 1930s, including one of a chick-embryo nerve fibre reaching out to find connections.

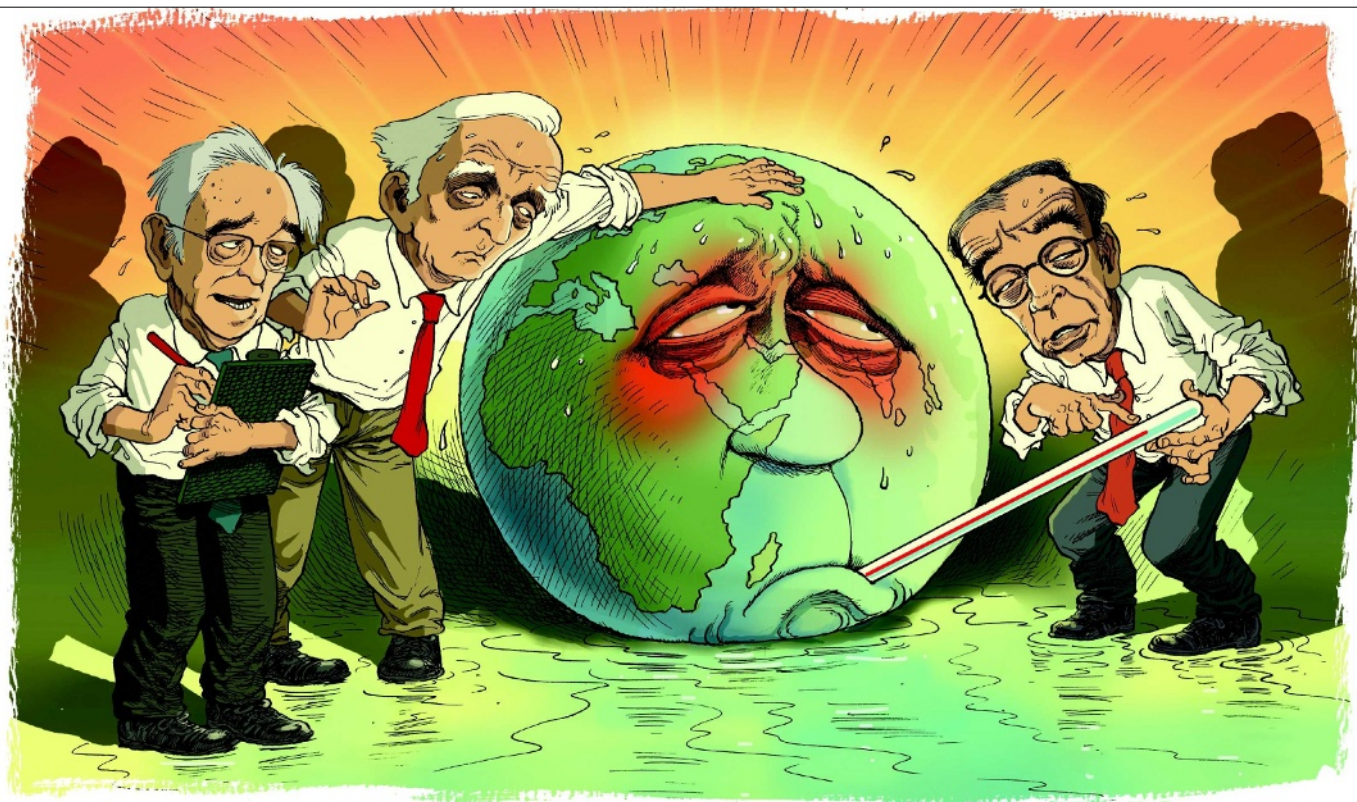
Giacomo left his skeleton to his beloved museum when he died in 1898. At its feet sits a bell jar containing his brain, preserved according to his own method. In an 1899 lecture at Turin's academy of medicine, a colleague described its anatomical peculiarities — including the presence of a very rare double Rolando sulcus. Lombroso, listening avidly in the audience, was thrilled. Proof, he cried, of Giacomo's genius and of Lombroso's own theory. ■

Alison Abbott is *Nature's* senior European correspondent.

For more Hidden treasures, see www.nature.com/nature/focus/hidden treasures.

MUS. HUM. ANATOMY, UNIV. TURIN

ESSAY



D. PARKINS

Madrid 1995: Diagnosing climate change

John Houghton chaired the tense IPCC meeting without which there would be no Kyoto Protocol. Here he recalls how science won the day.

Climate change is increasingly recognized as one of the most serious threats to humankind. But it was not always so. That the majority of world leaders are now convinced of this peril, even though its full effects will not be felt for decades, is due largely to the work of one organization: the Intergovernmental Panel on Climate Change (IPCC).

The IPCC draws together hundreds of the world's top scientists to study and assess the effects and consequences of human activity on the climate. As an intergovernmental organization under the United Nations, its five-yearly assessments must be accepted by governments, and the IPCC recommendations are then taken as the basis for policy. Crucially, the content of the final text is determined solely by the science, based on the research literature.

As first chair, and later co-chair, of the IPCC's working group on the physical basis of climate change from 1988 to 2002, I was present at all the meetings associated with the first three IPCC Reports — in 1990, 1995 and 2001. One meeting was particularly challenging, contentious and ultimately satisfying: this was to agree the IPCC's second assessment report, and was

held in Madrid over three days at the end of November 1995.

The mid-1990s were a crucial time in the study of climate change. The science had moved on since the first IPCC assessment in 1990 and there was growing evidence for the impact of human activity. Climate change had stirred worldwide interest, and groups whose interests might be affected by any policies derived from the report were keen to have their say. We knew that the process of agreeing a final text for the 1995 assessment was going to be more controversial than before.

Climate wars

The Madrid meeting was attended by 177 government delegates from 96 countries, 28 scientists representing the authors of the 11 chapters of the IPCC report, and 14 representatives of non-governmental organizations (NGOs). Some had strong environmentalist agendas. Others represented business and industry. One industry NGO, the Global Climate Coalition (now "deactivated"), was among the most active in challenging the IPCC's conclusions on climate change. Backed by powerful parts of the US and

international energy industry, the coalition had begun to take part in the IPCC's activities by attending meetings, and writing and reviewing papers. Its representatives worked closely with the official delegates from Saudi Arabia, Kuwait and other oil- and gas-exporting states.

In preparation for the meeting, a small group containing the leading scientist authors had prepared a shortened version of the main assessment report — a Summary for Policymakers. The aim of the meeting would be to shape this summary text into a form that all those present (government representatives and IPCC scientists) could agree on.

The meeting opened under the co-chairmanship of myself and Luiz Gylvan Meira Filho, former head of Brazil's space agency. We began by discussing the content and conclusions of the 11 chapters, and our deliberations did not seem out of the ordinary compared with previous meetings.

By the second day, as we got to grips with the Summary for Policymakers, things were moving much more slowly than expected. A small core of delegates, led by the representatives of Saudi Arabia and Kuwait, wanted to weaken the

statements about the extent of climate change and put greater emphasis on the uncertainties associated with the causes of climate change.

Inevitably, the most heated discussions focused on how to summarize Chapter 8 of the full report, which addressed the extent to which human-induced global warming had been observed in the climate record. Before the meeting, several governments, in particular the United States, had pointed out that the summary at the beginning of Chapter 8 was not entirely consistent with the conclusions at the end of the chapter, and that some statements had not been crafted with sufficient care. When the plenary session began, many other countries echoed these views and asked for revisions, which the meeting accepted.

It was during this debate that the delegates from Kuwait and Saudi Arabia asked for words or phrases that emphasized uncertainty regarding human-induced climate change to be inserted in the Summary for Policymakers. Two sentences in the summary were particularly sensitive. The first concerned patterns of temperature change that scientists were observing across the globe. In its draft form it read:

More convincing evidence for the attribution of a human effect on climate is emerging from pattern based studies, in which the modelled climate response to combined forcing by greenhouse gases and anthropogenic sulphate aerosols is compared with observed geographical, seasonal and vertical patterns of atmospheric temperature change."

Saudi Arabia and Kuwait requested that the first three words be altered to "Some preliminary evidence", because the draft version of Chapter 8 had used the phrase "preliminary evidence" in association with pattern-based studies. For an hour and a half, the meeting debated the appropriateness of "preliminary" instead of "convincing". All other delegates who spoke argued that the evidence in Chapter 8 did not warrant the description "preliminary", and a sentence was finally drafted that began "More convincing recent evidence...". The meeting agreed that a footnote should be added to the final text explaining that Saudi Arabia and Kuwait dissented from the majority view. This footnote was later withdrawn at the countries' own request.

The second sentence to spark disagreement read:

The balance of evidence suggests human influence on global climate.

In 1990, the IPCC's report had concluded that unequivocal detection of anthropogenic climate change "would not be likely for a decade or

more". So again Saudi Arabia swam against the tide, arguing that "balance of evidence" was far too strong. The debate over this sentence lasted well over an hour and various alternatives were proposed: "The balance/weight/preponderance of evidence points to/suggests/indicates an appreciable/detectable/measurable/discernible human influence on the global climate."

The sentence finally agreed on was: "The balance of evidence suggests a discernible human influence on global climate." The word that most excited the meeting was "discernible": it was suggested at a very late stage by the UK delegate as a replacement for "appreciable", and was spontaneously applauded.

It was a tough meeting but no one lost their cool, and the undoubted tension was punctuated by occasional humour. I joked with Mohamed Al-Sabban, the head of Saudi Arabia's delegation, that if I were a member of his delegation I would advise strongly against a particular proposal he was making, as being against his country's interests. He promptly made me an honorary citizen of Saudi Arabia, later referring a few times to my 'honorary citizenship'.

It all finished at 20 minutes past midnight, 10 minutes before the conference centre was due to close, and to my relief no unsatisfactory compromises had been made. When I read through what we had agreed, I realized that the powerful discipline of science had created a final result that was better in its clarity, accuracy and presentation than the draft text with which we had begun.

Integrity attack

But that is not how our critics interpreted things. The following year, when the full, revised IPCC assessment was published, the Global Climate Coalition accused the IPCC of dishonesty and fraud. On 12 June 1996, the *Wall Street Journal* published an op-ed by Frederick Seitz, a former president of the US National Academy of Sciences, that accused the IPCC authors of altering Chapter 8 of the 1995 report after it had been formally approved. "I have never witnessed a more disturbing corruption of the peer-review process than the events that led to this IPCC report," Seitz claimed, quoting differences between the draft chapter that had been available to the Madrid meeting and the final published version.

The late Bert Bolin (the IPCC chairman at the time), Gylvan Meira Filho and I wrote to the *Wall Street Journal* explaining that there was no question of 'corruption' of any kind and that the authors had been working completely within IPCC rules. The newspaper printed a much-shortened version of our reply on 25 June

but continued to run articles repeating Seitz's accusations. In addition, for some months afterwards the Global Climate Coalition waged a relentless campaign in the United States against Ben Santer, the lead author of Chapter 8 — to the extent that the American Meteorological Society felt compelled to publish an open letter praising him and supporting the IPCC.

Happily, the IPCC came through the process with its integrity intact. In 2001 and 2007 it released two further reports that substantially confirmed the conclusions of the first two reports, adding much more detail about the likely impacts of climate change and disclosing significant evidence about positive feedbacks amplifying global warming that had barely been recognized in 1995.

I am often asked to explain how it is possible for so many scientists from different countries and cultures to come to a consensus about a subject as complex and uncertain as

climate change. My reply is that science, with its emphasis on robust data, repeatability, balance, accuracy and integrity and its reliance on argument and debate to reach a conclusion, provides an ideal process for such an endeavour.

Furthermore, the consensus does not, of course, apply to everything: the IPCC has always been at pains to distinguish between what we know with reasonable certainty and those areas in which much is unclear.

Perhaps the most important outcome of the IPCC process has been the creation of a large research community, well informed about all aspects of climate change, which can provide sound information to reduce the confusion bound to arise from such a complex nexus of science and society. The involvement of this community in all aspects of the assessment process has meant that the IPCC's published reports are more punchy and more precise than they would otherwise have been.

In Madrid in 1995, the IPCC scientific assessment process, based on the findings of the latest research, was sorely tested. Had the science not come through unscathed, the integrity of the panel would have been seriously questioned, and governments would have faltered on taking urgent action on climate change, such as the signing in 1997 of the Kyoto Protocol.

John Houghton chaired the IPCC meeting in Madrid from 27 to 29 November 1995. He was chairman of the UK Royal Commission on Environmental Pollution from 1992 to 1996 and is currently president of the John Ray Initiative. e-mail: john.houghton@jri.org.uk

"The undoubted tension was punctuated by occasional humour."

For more Meetings that Changed the World, see www.nature.com/nature/focus/meetings.

NEWS & VIEWS

GERONTOLOGY

Healthy old age

Thomas B. L. Kirkwood

Do longer lives mean that growing numbers of us will spend more time in a state of high-cost dependency? Evidence from one elderly cohort suggests that excessive levels of disability are far from inevitable.

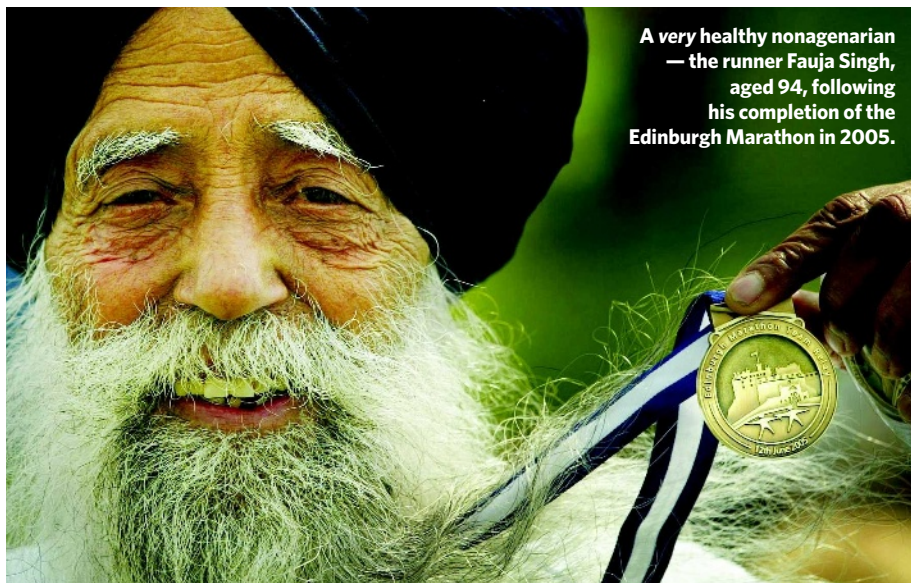
Birthdays are a sign of survival: the more you have, the longer you have lived. But survival is not the same thing as health. One of the major concerns, as life expectancies around the world continue to climb, is whether longer life means, on average, a longer period spent suffering from age-related frailty, disability and disease^{1,2}. A study from Denmark, just published by Christensen *et al.* in *Proceedings of the National Academy of Sciences*³, reassuringly shows that it is not necessarily so. Christensen *et al.* analysed longitudinal data from 2,262 individuals in the Danish population born in 1905 who were still alive in 1998 (about two-thirds of the surviving total), in order to find out whether living to an exceptional age brought with it a very high risk of disability and dependency.

To track how the nonagenarians fared, assessments were made four times between 1998 and 2005 (that is, between the ages of 92 and 100). The assessments included measures of the ability to carry out daily activities (such as shopping, cooking and bathing), cognitive performance, and disability and physical status. About half of the subjects were living in a house or apartment, one-third in a nursing home, and the rest were in intermediate-care environments.

When individuals were categorized as being independent or non-independent (unable to undertake basic activities without assistance, or having significantly impaired cognition), the proportion capable of independent living showed surprisingly little decline. Between the ages of 92 and 100, it fell only from 39% to 33%. Even centenarians had a one-third chance of remaining independent.

Statistics such as these need careful dissection. The number still alive in 1998 was more than ten times greater than the number who made it through to 2005. Statistically speaking, it is as hard to get from 92 to 100 as it is to reach 92 in the first place. Thus, the proportion that remains independent at age 100 is a fraction of a very much reduced denominator. To understand what is happening, therefore, it is essential to consider individual trajectories of health and disability.

When Christensen *et al.* followed individual



A very healthy nonagenarian
— the runner Fauja Singh,
aged 94, following
his completion of the
Edinburgh Marathon in 2005.

D. CHESKIN/PA

trajectories across the duration of their study, the data showed a progressive trend towards increased likelihood of non-independence. The fact that this was not so apparent in the cross-sectional statistics is easily explained. Those who lose independence first tend to die first. In effect, the numerically small group of individuals who survive to 100 were probably among the most independent subset of individuals at 92. Those who remained alive but were now dependent at 100 would have mostly been independent at 92. They did not, on the whole, endure nearly a decade in a state of dependency.

In terms of longevity, men do not fare as well as women. In the 1905 Danish cohort, women outnumbered men by about three to one in 1998 and by about four to one in 2005. For women, however, the fraction remaining independent showed a faster decline, from 37% to 28%. The fraction remaining independent among the men was higher at all ages across this time interval, beginning at 45% in 1998 and actually finishing higher at 53% in 2005. Although women live longer than men, it is well known that they experience a longer period of age-related disability and ill health⁴. The Danish data seem to bear this out, although

for both women and men the evidence clearly signals that a longer life does not in itself mean an extended period of dependency.

There is much doom and gloom globally about the costs of population ageing, in terms both of costs to society through the provision of medical and social support to a growing fraction of dependent older people, and of the price paid by the individual in lost quality of life. The study by Christensen *et al.*³ therefore provides a welcome ray of light. It is of course limited by focusing on only a single birth cohort in a single country, where social and medical services are of an unusually high standard, and by providing information only about those who were already somewhat unusual in having survived to 92 at the inception of the research. Nevertheless, it shows that the oldest old do not, merely by living longer, make a dramatically increased contribution to the 'burden' of dependency. Indeed, it is a widespread but gross misperception that population ageing is the main driver of increased costs of care⁵. The truth is that it is care during the terminal months of life that is expensive. This is so whether the terminal illness occurs early or late. The fact that most terminal care is incurred by people who are

old simply reflects the fact that most of us now reach old age.

For the future, it is essential that strong links are forged between demographic studies, such as that on the Danish cohort born in 1905, and basic research on the biology of ageing. The challenge is clear. How can we better understand the underlying mechanisms of ageing so as to enhance further the quality of the later years of life? To deliver this outcome, we need to understand much more about the factors that influence longevity and their relationships with age-related disability and disease⁶. Ageing is more intrinsically malleable than had been thought⁷, and we need to exploit this discovery to increase the individual health span as a

fraction of the individual lifespan, however long the latter may be.

Thomas B. L. Kirkwood is at the Institute for Ageing and Health, Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne NE4 5PL, UK.

e-mail: tom.kirkwood@ncl.ac.uk

1. Crimmins, E. M. *Annu. Rev. Public Health* **25**, 79–98 (2004).
2. Manton, K. G. *Annu. Rev. Public Health* **29**, 91–113 (2008).
3. Christensen, K., McGue, M., Petersen, L., Jeune, B. & Vaupel, J. W. *Proc. Natl Acad. Sci. USA* **105**, 13274–13279 (2008).
4. Jagger, C. et al. *J. Gerontol. Med. Sci.* **62A**, 408–414 (2007).
5. Lubitz, J. et al. *N. Engl. J. Med.* **349**, 1048–1055 (2003).
6. Olshansky, S. J., Perry, D., Miller, R. A. & Butler, R. N. *Ann. NY Acad. Sci.* **1114**, 11–13 (2007).
7. Kirkwood, T. B. L. *Cell* **120**, 437–447 (2005).

MOLECULAR BIOLOGY

DNA endgames

Hannah L. Klein

If it is not repaired efficiently, damage to DNA double strands can have dire consequences for both the cell and the organism. Given the gravity of this situation, cells use two pathways to start the process.

In human cells, the two strands of DNA that make up each of our chromosomes break several times a day, either during normal cellular metabolism or following exposure to DNA-damaging agents. Efficient repair of such damage is crucial for restoring the integrity of the genome, and also for preventing premature termination of a process that occurs during meiotic cell division to ensure the genetic diversity of our progeny. Studies in bacteria have shed light on how double-strand breaks are processed for mending, but uncertainty remains over many pivotal aspects of DNA repair in eukaryotic organisms (yeast, plants and animals). Three studies^{1–3}, including one by Mimitou and Symington¹ on page 770 of this issue, unlock some of the mysteries of the initiating steps in eukaryotic DNA repair.

Whether during replication or repair, if it is to extend, a DNA strand needs a complementary sequence to use as a template. In the case of DNA double-strand breaks (DSBs), the complementary copies of DNA are severed in the same place and so cannot act as templates for each other. Such breaks are repaired through two mechanisms — non-homologous end-joining and homologous recombination. Of the two, homologous recombination is by far the more accurate.

The two ends of a DNA strand are termed the 3' and 5' ends, and it is thought that homologous recombination starts with the nucleolytic

degradation of DSB ends in the 5' to 3' direction to produce 3' single-stranded DNA 'tails'. These can be used for repair by recruiting an intact complementary sequence as a template. Indeed, in yeast, intermediate sequences that have 3' single-stranded tails,

several hundred nucleotides long, are formed from processing the broken ends of the double strands^{4,5}. But what molecules and pathways mediate DSB processing during homologous recombination?

One potential candidate is the Mre11 complex, which is recruited to the site of DSBs soon after damage and is known to be involved in processing the broken ends. The Mre11 protein, which forms part of this complex, has nuclease enzymatic activity (that is, it can split the bonds between nucleotides), which is essential for homologous recombination^{6,7}; but it catalyses degradation only in the 3' to 5' direction, not in the required 5' to 3' direction⁸. A protein that interacts with the Mre11 complex, Sae2/CtIP, also acts early in the processing of DSB ends⁹ and has nuclease activity¹⁰; however, it does not extensively degrade DSB ends. Finally, another nuclease, Exo1, is involved in cell division, particularly in processing the ends created during DNA replication. But genetic studies have been unable to confirm that Exo1 is the end-processing nuclease in homologous recombination, and so the search has continued.

Mimitou and Symington¹ now describe an alternative pathway for end processing that acts without Exo1. The authors were inspired by studies demonstrating¹¹ that, in bacteria, DSB ends are processed by several mechanisms and protein factors. For example, in one end-processing pathway, *Escherichia coli* uses the helicase enzyme RecQ, together with the RecJ nuclease, acts at an early stage of homologous recombination to initiate nucleolytic degradation; at a later stage, RecQ alone processes the intermediates further to complete the repair process.

The yeast *Saccharomyces cerevisiae* has one RecQ-related protein — Sgs1. Mutations in the gene encoding Sgs1 lead to defects in homologous recombination, but do not completely abolish end processing or meiotic recombination. Therefore, Sgs1 was not considered to have an early role in homologous recombination at the end-processing stage. However, *sgs1* mutants show gross chromosomal rearrangements, and, at the rearrangement breakpoints, the two DNA strands are not entirely complementary¹². Such rearrangements could result from either an early defect in end processing or a late defect in preventing recombination events. Nonetheless, previous studies had suggested that Sgs1 and related eukaryotic RecQ helicases act only at a late stage in recombination to prevent such types of rearrangement.

Mimitou and Symington¹, and two other groups^{2,3}, now show in *S. cerevisiae* that, together with Sae2, Sgs1 acts early in homologous recombination, in end processing, and that this Sgs1 activity occurs in parallel with that

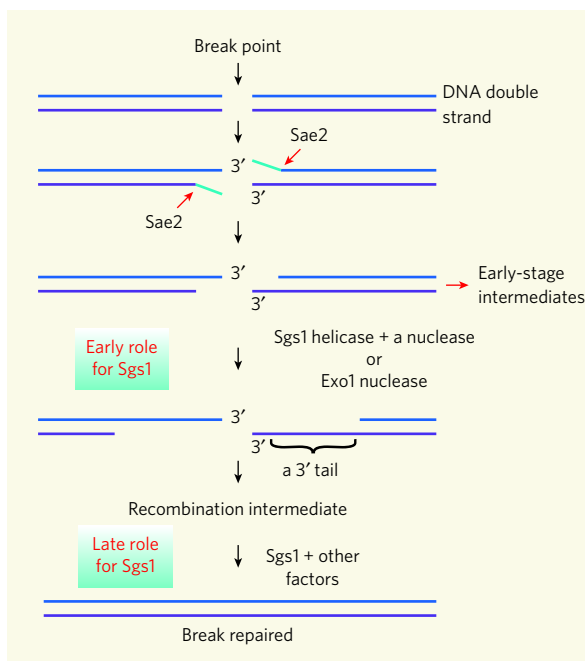


Figure 1 | Processing DNA's broken ends. New studies^{1–3} show that end processing after damage to DNA double strands proceeds in two steps. First, Sae2, together with the Mre11 complex (not shown), prepares the ends for degradation by trimming them. Next, the ends are resected, either by Sgs1 plus a nuclease or by the Exo1 nuclease, to give 3' tails. Formation of recombinant intermediates and break repair follow.

old simply reflects the fact that most of us now reach old age.

For the future, it is essential that strong links are forged between demographic studies, such as that on the Danish cohort born in 1905, and basic research on the biology of ageing. The challenge is clear. How can we better understand the underlying mechanisms of ageing so as to enhance further the quality of the later years of life? To deliver this outcome, we need to understand much more about the factors that influence longevity and their relationships with age-related disability and disease⁶. Ageing is more intrinsically malleable than had been thought⁷, and we need to exploit this discovery to increase the individual health span as a

fraction of the individual lifespan, however long the latter may be.

Thomas B. L. Kirkwood is at the Institute for Ageing and Health, Campus for Ageing and Vitality, Newcastle University, Newcastle upon Tyne NE4 5PL, UK.

e-mail: tom.kirkwood@ncl.ac.uk

1. Crimmins, E. M. *Annu. Rev. Public Health* **25**, 79–98 (2004).
2. Manton, K. G. *Annu. Rev. Public Health* **29**, 91–113 (2008).
3. Christensen, K., McGue, M., Petersen, I., Jeune, B. & Vaupel, J. W. *Proc. Natl Acad. Sci. USA* **105**, 13274–13279 (2008).
4. Jagger, C. et al. *J. Gerontol. Med. Sci.* **62A**, 408–414 (2007).
5. Lubitz, J. et al. *N. Engl. J. Med.* **349**, 1048–1055 (2003).
6. Olshansky, S. J., Perry, D., Miller, R. A. & Butler, R. N. *Ann. NY Acad. Sci.* **1114**, 11–13 (2007).
7. Kirkwood, T. B. L. *Cell* **120**, 437–447 (2005).

MOLECULAR BIOLOGY

DNA endgames

Hannah L. Klein

If it is not repaired efficiently, damage to DNA double strands can have dire consequences for both the cell and the organism. Given the gravity of this situation, cells use two pathways to start the process.

In human cells, the two strands of DNA that make up each of our chromosomes break several times a day, either during normal cellular metabolism or following exposure to DNA-damaging agents. Efficient repair of such damage is crucial for restoring the integrity of the genome, and also for preventing premature termination of a process that occurs during meiotic cell division to ensure the genetic diversity of our progeny. Studies in bacteria have shed light on how double-strand breaks are processed for mending, but uncertainty remains over many pivotal aspects of DNA repair in eukaryotic organisms (yeast, plants and animals). Three studies^{1–3}, including one by Mimitou and Symington¹ on page 770 of this issue, unlock some of the mysteries of the initiating steps in eukaryotic DNA repair.

Whether during replication or repair, if it is to extend, a DNA strand needs a complementary sequence to use as a template. In the case of DNA double-strand breaks (DSBs), the complementary copies of DNA are severed in the same place and so cannot act as templates for each other. Such breaks are repaired through two mechanisms — non-homologous end-joining and homologous recombination. Of the two, homologous recombination is by far the more accurate.

The two ends of a DNA strand are termed the 3' and 5' ends, and it is thought that homologous recombination starts with the nucleolytic

degradation of DSB ends in the 5' to 3' direction to produce 3' single-stranded DNA 'tails'. These can be used for repair by recruiting an intact complementary sequence as a template. Indeed, in yeast, intermediate sequences that have 3' single-stranded tails,

several hundred nucleotides long, are formed from processing the broken ends of the double strands^{4,5}. But what molecules and pathways mediate DSB processing during homologous recombination?

One potential candidate is the Mre11 complex, which is recruited to the site of DSBs soon after damage and is known to be involved in processing the broken ends. The Mre11 protein, which forms part of this complex, has nuclease enzymatic activity (that is, it can split the bonds between nucleotides), which is essential for homologous recombination^{6,7}; but it catalyses degradation only in the 3' to 5' direction, not in the required 5' to 3' direction⁸. A protein that interacts with the Mre11 complex, Sae2/CtIP, also acts early in the processing of DSB ends⁹ and has nuclease activity¹⁰; however, it does not extensively degrade DSB ends. Finally, another nuclease, Exo1, is involved in cell division, particularly in processing the ends created during DNA replication. But genetic studies have been unable to confirm that Exo1 is the end-processing nuclease in homologous recombination, and so the search has continued.

Mimitou and Symington¹ now describe an alternative pathway for end processing that acts without Exo1. The authors were inspired by studies demonstrating¹¹ that, in bacteria, DSB ends are processed by several mechanisms and protein factors. For example, in one end-processing pathway, *Escherichia coli* uses the helicase enzyme RecQ, together with the RecJ nuclease, acts at an early stage of homologous recombination to initiate nucleolytic degradation; at a later stage, RecQ alone processes the intermediates further to complete the repair process.

The yeast *Saccharomyces cerevisiae* has one RecQ-related protein — Sgs1. Mutations in the gene encoding Sgs1 lead to defects in homologous recombination, but do not completely abolish end processing or meiotic recombination. Therefore, Sgs1 was not considered to have an early role in homologous recombination at the end-processing stage. However, *sgs1* mutants show gross chromosomal rearrangements, and, at the rearrangement breakpoints, the two DNA strands are not entirely complementary¹². Such rearrangements could result from either an early defect in end processing or a late defect in preventing recombination events. Nonetheless, previous studies had suggested that Sgs1 and related eukaryotic RecQ helicases act only at a late stage in recombination to prevent such types of rearrangement.

Mimitou and Symington¹, and two other groups^{2,3}, now show in *S. cerevisiae* that, together with Sae2, Sgs1 acts early in homologous recombination, in end processing, and that this Sgs1 activity occurs in parallel with that

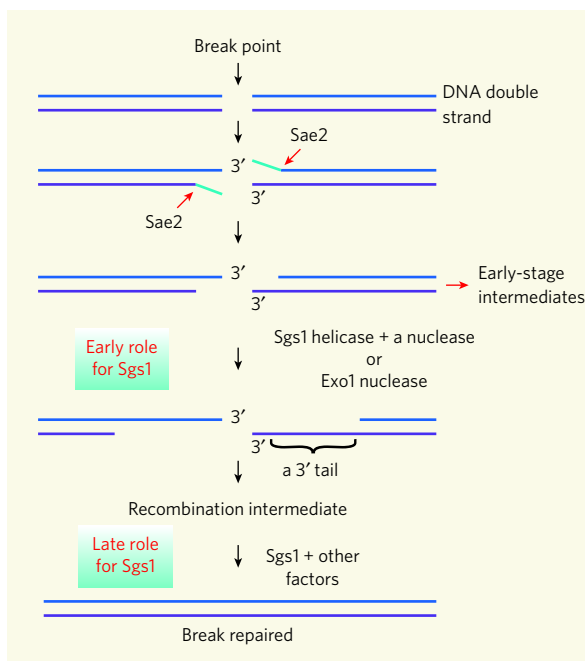


Figure 1 | Processing DNA's broken ends. New studies^{1–3} show that end processing after damage to DNA double strands proceeds in two steps. First, Sae2, together with the Mre11 complex (not shown), prepares the ends for degradation by trimming them. Next, the ends are resected, either by Sgs1 plus a nuclease or by the Exo1 nuclease, to give 3' tails. Formation of recombinant intermediates and break repair follow.

of the Exo1 nuclease. Using DSB substrates generated *in vivo* that require lengthy resection for repair, the authors followed the processing of the ends under normal conditions, or when Sae2, Sgs1 and/or Exo1 were absent.

They find that Sae2, in complex with Mre11 and its partners, initiates end processing to form the first intermediates. Next, either Exo1 or Sgs1 plus another nuclease — possibly Dna2 — resect the ends to give the required 3' tails (Fig. 1). When both of these second-stage end-processing pathways are defective and only Sae2 is active, a ladder of fragments, corresponding to incremental cleavages of 50–100 nucleotides, is seen, and end processing and homologous recombination are in general as defective as in the most severe mutant forms of homologous recombination. Thus, processing by Sae2, and by Mre11 and its associates, prepares the DSB ends for subsequent resection by either Exo1 or Sgs1.

The implications of these findings are wide-ranging. When mutated, the human RecQ helicases BLM, WRN and RECQL4 (and possibly RECQL5) have been implicated in genomic instability, a predisposition to cancer, and ageing. The defects caused by these mutations are thought to result from an inability to control the outcome of a recombination event during the late stages of homologous recombination. Consequently, inappropriate recombination events occur, and failure to resolve the intermediates properly leads to disease-associated mutations referred to as loss of heterozygosity. In light of the latest findings^{1–3}, we must consider the possibility that defective end processing, even if it is normally mediated by several pathways, could lead to incomplete repair of a DSB, with little regard for complementarity between the two strands and so leading to mutations and rearrangements.

As with all major advances, the findings^{1–3} raise more questions than they answer. The studies focused on recombination during mitotic cell division — the basic cell-division cycle that produces most cells in the body. Does regulation of end processing differ between meiosis and mitosis? What controls which resection pathway — that mediated by Exo1 or by Sgs1 — is used? Is it decided according to whether the break is caused by a cellular process or external DNA damage, the context of the damage, or the stage of the cell cycle at which it occurs?

And lastly, what are the effects of mutations in human EXO1? This nuclease would be expected to have a crucial role in repairing DSB damage. But unlike human RecQ helicases, EXO1 has not been linked to disease. It is possible that, in humans, the RecQ helicases have a more prominent role in DSB end processing than EXO1 and so the effects of mutations in the gene encoding this latter nuclease are more subtle. EXO1 mutations might be one of many factors that increase the predisposition to cancer or lead to infertility, which is associated with abnormal meiosis. The relative contributions of

the human RecQ helicases and EXO1 to DSB end processing are still to be defined.

Hannah L. Klein is in the Department of Biochemistry, New York University School of Medicine, 550 First Avenue, New York, New York 10016, USA.

e-mail: hannah.klein@nyumc.org

1. Mimitou, E. P. & Symington, L. S. *Nature* **455**, 770–774 (2008).
2. Zhu, Z., Chung, W.-H., Shim, E. Y., Lee, S. E. & Ira, G. *Cell* **134**, 981–994 (2008).

3. Gravel, S., Chapman, J. R., Magill, C. & Jackson, S. P. *Genes Dev.* **22**, 2767–2772 (2008).
4. Sun, H., Treco, D. & Szostak, J. W. *Cell* **64**, 1155–1161 (1991).
5. White, C. I. & Haber, J. E. *EMBO J.* **9**, 663–673 (1990).
6. Furuse, M. *et al.* *EMBO J.* **17**, 6412–6425 (1998).
7. Llorente, B. & Symington, L. S. *Mol. Cell. Biol.* **24**, 9682–9694 (2004).
8. Paull, T. T. & Gellert, M. *Mol. Cell* **1**, 969–979 (1998).
9. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. *J. Biol. Chem.* **280**, 38631–38638 (2005).
10. Lengsfeld, B. M. *et al.* *Mol. Cell* **28**, 638–651 (2007).
11. Amundsen, S. K. & Smith, G. R. *Cell* **112**, 741–744 (2003).
12. Myung, K., Datta, A., Chen, C. & Kolodner, R. D. *Nature Genet.* **27**, 113–116 (2001).

SOLID-STATE PHYSICS

Recipe for spin currents

N. P. Ong

Generating currents that rely on the spins of electrons to make electronic devices requiring less power is both desirable and daunting. A neat way of creating such currents eases that task.

The field of spintronics — spin electronics — seeks to harness the spin of electrons in metals and semiconductors in order to perform tasks that are, at present, routinely carried out by electrons' charge. Spintronics offers a promising path to achieve further reduction in both the size and power consumption of solid-state devices. The past few years have seen many ingenious experiments¹ clarifying the basic principles that govern the creation and manipulation of spin currents — the flow of electrons with a net spin polarization. On page 778 of this issue, Uchida *et al.*² introduce the 'spin Seebeck effect' as an especially elegant way of generating spin currents from a spin-voltage source that can be maintained across a distance of roughly 1 cm at room temperature.

The ordinary Seebeck effect refers to the appearance of a voltage when the ends of an electrically insulated wire are held at different temperatures (Fig. 1a, overleaf). The temperature gradient drives a heat current that is carried by electrons to the cooler end. Because the wire is electrically insulated, there is a backflow of electrons driven by a finite gradient in the chemical potential, μ — the energy of the most energetic electrons in the wire. The situation is analogous to using a fan to pile up water at one end of a fish tank. The current of surface water flowing downwind is balanced by a backflow driven by the pressure gradient in the water. At steady state, the water level in the tank (the analogue of μ) has a tilted profile. Likewise, μ exhibits a tilted profile in a temperature gradient (Fig. 1a), and the voltage observed is the difference in μ at the two ends. This thermoelectric voltage can be used to generate electricity, and is a vital source of energy in space satellites. The reverse effect, known as the Peltier effect, is used to cool materials ranging from microchips to rare wines.

The experiment carried out by Uchida *et al.*² extends the Seebeck effect to spins. In a ferromagnet — in contrast to non-magnetic metals, such as copper and platinum — the density of spin-up electrons greatly exceeds that of spin-down electrons at thermal equilibrium. A key point in the field of spintronics is that, in a non-equilibrium situation, the chemical potential of the spin-up electrons, μ_{\uparrow} , can be made to deviate substantially from that of spin-down electrons, μ_{\downarrow} . The different gradients in μ_{\uparrow} and μ_{\downarrow} then lead to distinct spin-up and spin-down currents flowing in the metal. The difference in chemical potential, $\mu_{\uparrow} - \mu_{\downarrow}$, is called spin voltage.

Several experiments¹ have investigated spin currents driven by a spin voltage that is maintained across relatively short distances. For example, when an electrical current flows from a ferromagnet into copper, both the spin-up excess and spin voltage persist in the copper over about 500 nm before vanishing. In their experiment, Uchida and colleagues obtained a spin voltage that extends over a macroscopic distance of 6 mm. They applied a temperature gradient to a thin film of a nickel–iron ferromagnet, and showed that μ_{\uparrow} and μ_{\downarrow} display distinct tilted profiles. This results in more spin-up electrons accumulating at, say, the cooler end, and more spin-down electrons at the warmer end. Hence, the spin voltage has opposite signs at the two ends (Fig. 1b).

To prove that this is indeed the case, Uchida *et al.* allowed the spin voltage to drive a spin current into a platinum film deposited across one end of the nickel–iron film. As shown in Figure 1c, the temperature gradient enables more spin-up than spin-down electrons to accumulate at the cool end of the ferromagnet (spin-up is here defined to be in the direction of the applied magnetic field). In turn, the higher μ_{\uparrow} implied by the excess spin-up population

of the Exo1 nuclease. Using DSB substrates generated *in vivo* that require lengthy resection for repair, the authors followed the processing of the ends under normal conditions, or when Sae2, Sgs1 and/or Exo1 were absent.

They find that Sae2, in complex with Mre11 and its partners, initiates end processing to form the first intermediates. Next, either Exo1 or Sgs1 plus another nuclease — possibly Dna2 — resect the ends to give the required 3' tails (Fig. 1). When both of these second-stage end-processing pathways are defective and only Sae2 is active, a ladder of fragments, corresponding to incremental cleavages of 50–100 nucleotides, is seen, and end processing and homologous recombination are in general as defective as in the most severe mutant forms of homologous recombination. Thus, processing by Sae2, and by Mre11 and its associates, prepares the DSB ends for subsequent resection by either Exo1 or Sgs1.

The implications of these findings are wide-ranging. When mutated, the human RecQ helicases BLM, WRN and RECQL4 (and possibly RECQL5) have been implicated in genomic instability, a predisposition to cancer, and ageing. The defects caused by these mutations are thought to result from an inability to control the outcome of a recombination event during the late stages of homologous recombination. Consequently, inappropriate recombination events occur, and failure to resolve the intermediates properly leads to disease-associated mutations referred to as loss of heterozygosity. In light of the latest findings^{1–3}, we must consider the possibility that defective end processing, even if it is normally mediated by several pathways, could lead to incomplete repair of a DSB, with little regard for complementarity between the two strands and so leading to mutations and rearrangements.

As with all major advances, the findings^{1–3} raise more questions than they answer. The studies focused on recombination during mitotic cell division — the basic cell-division cycle that produces most cells in the body. Does regulation of end processing differ between meiosis and mitosis? What controls which resection pathway — that mediated by Exo1 or by Sgs1 — is used? Is it decided according to whether the break is caused by a cellular process or external DNA damage, the context of the damage, or the stage of the cell cycle at which it occurs?

And lastly, what are the effects of mutations in human EXO1? This nuclease would be expected to have a crucial role in repairing DSB damage. But unlike human RecQ helicases, EXO1 has not been linked to disease. It is possible that, in humans, the RecQ helicases have a more prominent role in DSB end processing than EXO1 and so the effects of mutations in the gene encoding this latter nuclease are more subtle. EXO1 mutations might be one of many factors that increase the predisposition to cancer or lead to infertility, which is associated with abnormal meiosis. The relative contributions of

the human RecQ helicases and EXO1 to DSB end processing are still to be defined.

Hannah L. Klein is in the Department of Biochemistry, New York University School of Medicine, 550 First Avenue, New York, New York 10016, USA.

e-mail: hannah.klein@nyumc.org

1. Mimitou, E. P. & Symington, L. S. *Nature* **455**, 770–774 (2008).
2. Zhu, Z., Chung, W.-H., Shim, E. Y., Lee, S. E. & Ira, G. *Cell* **134**, 981–994 (2008).

3. Gravel, S., Chapman, J. R., Magill, C. & Jackson, S. P. *Genes Dev.* **22**, 2767–2772 (2008).
4. Sun, H., Treco, D. & Szostak, J. W. *Cell* **64**, 1155–1161 (1991).
5. White, C. I. & Haber, J. E. *EMBO J.* **9**, 663–673 (1990).
6. Furuse, M. *et al.* *EMBO J.* **17**, 6412–6425 (1998).
7. Llorente, B. & Symington, L. S. *Mol. Cell. Biol.* **24**, 9682–9694 (2004).
8. Paull, T. T. & Gellert, M. *Mol. Cell* **1**, 969–979 (1998).
9. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. *J. Biol. Chem.* **280**, 38631–38638 (2005).
10. Lengsfeld, B. M. *et al.* *Mol. Cell* **28**, 638–651 (2007).
11. Amundsen, S. K. & Smith, G. R. *Cell* **112**, 741–744 (2003).
12. Myung, K., Datta, A., Chen, C. & Kolodner, R. D. *Nature Genet.* **27**, 113–116 (2001).

SOLID-STATE PHYSICS

Recipe for spin currents

N. P. Ong

Generating currents that rely on the spins of electrons to make electronic devices requiring less power is both desirable and daunting. A neat way of creating such currents eases that task.

The field of spintronics — spin electronics — seeks to harness the spin of electrons in metals and semiconductors in order to perform tasks that are, at present, routinely carried out by electrons' charge. Spintronics offers a promising path to achieve further reduction in both the size and power consumption of solid-state devices. The past few years have seen many ingenious experiments¹ clarifying the basic principles that govern the creation and manipulation of spin currents — the flow of electrons with a net spin polarization. On page 778 of this issue, Uchida *et al.*² introduce the 'spin Seebeck effect' as an especially elegant way of generating spin currents from a spin-voltage source that can be maintained across a distance of roughly 1 cm at room temperature.

The ordinary Seebeck effect refers to the appearance of a voltage when the ends of an electrically insulated wire are held at different temperatures (Fig. 1a, overleaf). The temperature gradient drives a heat current that is carried by electrons to the cooler end. Because the wire is electrically insulated, there is a backflow of electrons driven by a finite gradient in the chemical potential, μ — the energy of the most energetic electrons in the wire. The situation is analogous to using a fan to pile up water at one end of a fish tank. The current of surface water flowing downwind is balanced by a backflow driven by the pressure gradient in the water. At steady state, the water level in the tank (the analogue of μ) has a tilted profile. Likewise, μ exhibits a tilted profile in a temperature gradient (Fig. 1a), and the voltage observed is the difference in μ at the two ends. This thermoelectric voltage can be used to generate electricity, and is a vital source of energy in space satellites. The reverse effect, known as the Peltier effect, is used to cool materials ranging from microchips to rare wines.

The experiment carried out by Uchida *et al.*² extends the Seebeck effect to spins. In a ferromagnet — in contrast to non-magnetic metals, such as copper and platinum — the density of spin-up electrons greatly exceeds that of spin-down electrons at thermal equilibrium. A key point in the field of spintronics is that, in a non-equilibrium situation, the chemical potential of the spin-up electrons, μ_{\uparrow} , can be made to deviate substantially from that of spin-down electrons, μ_{\downarrow} . The different gradients in μ_{\uparrow} and μ_{\downarrow} then lead to distinct spin-up and spin-down currents flowing in the metal. The difference in chemical potential, $\mu_{\uparrow} - \mu_{\downarrow}$, is called spin voltage.

Several experiments¹ have investigated spin currents driven by a spin voltage that is maintained across relatively short distances. For example, when an electrical current flows from a ferromagnet into copper, both the spin-up excess and spin voltage persist in the copper over about 500 nm before vanishing. In their experiment, Uchida and colleagues obtained a spin voltage that extends over a macroscopic distance of 6 mm. They applied a temperature gradient to a thin film of a nickel–iron ferromagnet, and showed that μ_{\uparrow} and μ_{\downarrow} display distinct tilted profiles. This results in more spin-up electrons accumulating at, say, the cooler end, and more spin-down electrons at the warmer end. Hence, the spin voltage has opposite signs at the two ends (Fig. 1b).

To prove that this is indeed the case, Uchida *et al.* allowed the spin voltage to drive a spin current into a platinum film deposited across one end of the nickel–iron film. As shown in Figure 1c, the temperature gradient enables more spin-up than spin-down electrons to accumulate at the cool end of the ferromagnet (spin-up is here defined to be in the direction of the applied magnetic field). In turn, the higher μ_{\uparrow} implied by the excess spin-up population

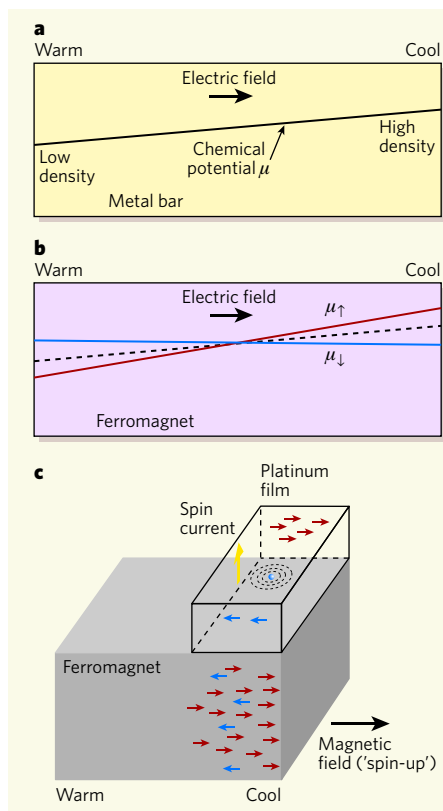


Figure 1 | The spin Seebeck effect. **a**, In the ordinary Seebeck effect, a temperature gradient in a metal bar causes more electrons to accumulate at the cool end, producing a tilt in the chemical potential (μ), which is observable as an electric field. **b**, Uchida *et al.*² extend the Seebeck effect to spins. In a ferromagnet, the temperature gradient results in an excess of spin-up electrons at the cool end, and an excess of spin-down electrons at the warm end. Their respective spin-chemical potentials, μ_{\uparrow} and μ_{\downarrow} , have tilt profiles of opposite signs (solid lines), the average (dashed line) giving the electric field. **c**, The spin Hall effect. The excess of spin-up electrons (red arrows) at the cool end of the ferromagnet drives a spin current that flows vertically into the platinum film (yellow arrow). Here, spin-up means that the direction of the spin is parallel to the magnetic field, and thus points to the right. By spin-orbit coupling, electrons 'see' a weak magnetic field circulating around a charged impurity (circles around blue dot). Scattering from the charged impurity causes spin-up electrons to accumulate preferentially on the far face of the platinum film, whereas spin-down electrons (blue arrows) end up on the near face. The imbalance is observed as a Hall voltage difference between the two faces.

drives a spin current into the platinum film. The decay of this spin current leads to the appearance of an electrical signal through the spin Hall effect^{3,4}, which enables spin currents to be detected using a sensitive voltmeter.

To understand the spin Hall effect, one can track an electron as it enters the platinum film and scatters off a charged impurity (Fig. 1c). The spin-orbit interaction — the interaction between an electron's spin and its motion — imparts a left-right asymmetry to the scattering

process. In the rest frame of the electron, the charged impurity rushing towards it constitutes a current filament, so the electron 'sees' a weak magnetic field circling the filament. This non-uniform magnetic field imparts a force on the electron along a direction that depends on its spin orientation^{3,4}. The net result is that spin-up electrons are pushed to the right of the impurity whereas spin-down electrons are pushed to its left.

In effect, each impurity acts like a spin filter that selectively kicks electrons to one side or the other, depending on their spin. As shown in Figure 1c, the excess spin-up population in the incident beam results in more charge accumulating on the far face than on the near face of the platinum film. The voltage difference between the two faces is observable as a Hall signal. The asymmetric scattering of electrons is especially large in materials with a high atomic number, such as platinum. Following its prediction^{3,4}, the spin Hall effect was first observed by applying purely optical techniques to semiconductors^{5,6}, and was later detected electrically in metals^{7,8}.

In a series of tests, Uchida *et al.*² convincingly show that the Hall voltage in the platinum film arises from the spin voltage. The Hall signal in

the platinum film tracks both the magnitude and the direction of the magnetization in the nickel-iron film. Moreover, by moving the platinum film along the length of the nickel-iron film, they show that the spin voltage varies linearly over the 6-mm length of the sample. In demonstrating that the spin Seebeck effect can produce a large, calibrated spin-voltage source that can be 'tapped' anywhere along the length of the ferromagnet, Uchida and colleagues have added an important tool to the spintronics toolbox. ■

N. P. Ong is in the Department of Physics, Princeton University, Princeton, New Jersey 08544, USA.

e-mail: npo@princeton.edu

1. Gregg, J. F. in *Spin Electronics* (eds Ziese, M. & Thornton, M. J.) 3–31 (Springer, 2001).
2. Uchida, K. *et al.* *Nature* **455**, 778–781 (2008).
3. D'yakonov, M. I. & Perel, V. I. *Phys. Lett. A* **35**, 459–460 (1971).
4. Hirsch, J. E. *Phys. Rev. Lett.* **83**, 1834–1837 (1999).
5. Kato, Y. K., Myers, R. C., Gossard, A. C. & Awschalom, D. D. *Science* **306**, 1910–1913 (2004).
6. Wunderlich, J., Kaestner, B., Sinova, J. & Jungwirth, T. *Phys. Rev. Lett.* **94**, 047204 (2005).
7. Valenzuela, S. O. & Tinkham, M. *Nature* **442**, 176–179 (2006).
8. Kimura, T., Otani, Y., Sato, T., Takahashi, S. & Maekawa, S. *Phys. Rev. Lett.* **98**, 156601 (2007).

NEUROSCIENCE

Brain's defence against cocaine

L. Judson Chandler and Peter W. Kalivas

Long-term exposure to cocaine changes the organization of synaptic connections within the addiction circuitry of the brain. This process might protect against the development and persistence of addiction.

Neurons modify their structure and communication with other neurons in response to experiences. Such experience-dependent neuroplasticity is crucial for survival because it allows learning from, and responses to, changes in the environment. But the cellular mechanisms that mediate this process can also be co-opted by drugs of abuse. Reporting in *Neuron*, Pulipparacharuvil *et al.*¹ describe how some of the chemical, structural and behavioural changes in neurons that are induced by repeated exposure to cocaine are regulated at a molecular level.

Drug addiction is characterized by compulsive drug seeking. It resembles a chronic relapsing disorder in which the addict resumes taking drugs after a period or periods of abstinence. Human and animal studies indicate that the recalcitrant nature of addiction results from drug-induced stimulation of reward-related learning processes in the brain. The pleasure-producing effects of the drug trigger cellular and molecular processes that are normally activated by natural rewards such as food and sex. Repeated exposure to an addictive drug

leads to a long-lasting associative memory of its rewarding properties through experience-dependent neuroplasticity. In effect, drug-seeking behaviour becomes hard-wired in the addict's brain, and the persistent memory trace is easily reactivated by drug-associated environmental stimuli, such as the sight of drug paraphernalia.

Along the dendritic processes of a neuron, morphologically specialized structures called dendritic spines receive most of the excitatory signals from other neurons through synaptic junctions. These spines are considered to be a primary cellular site for mediating the synaptic plasticity that is thought to underpin memory formation². One regulator of the density of excitatory signals on dendritic spines is the gene transcription factor MEF2 (ref. 3). When active (dephosphorylated), MEF2 favours elimination of dendritic spines, and when inactive (phosphorylated) it allows spine formation^{3,4}.

Repeated exposure to cocaine and other psychostimulants increases the number of dendritic spines on medium spiny neurons

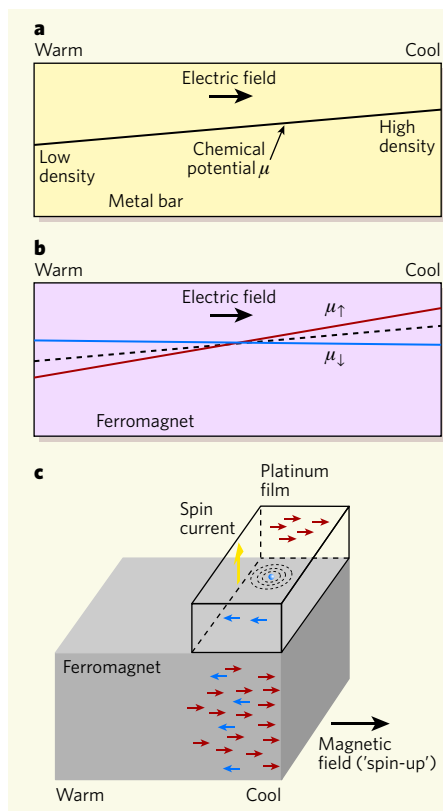


Figure 1 | The spin Seebeck effect. **a**, In the ordinary Seebeck effect, a temperature gradient in a metal bar causes more electrons to accumulate at the cool end, producing a tilt in the chemical potential (μ), which is observable as an electric field. **b**, Uchida *et al.*² extend the Seebeck effect to spins. In a ferromagnet, the temperature gradient results in an excess of spin-up electrons at the cool end, and an excess of spin-down electrons at the warm end. Their respective spin-chemical potentials, μ_{\uparrow} and μ_{\downarrow} , have tilt profiles of opposite signs (solid lines), the average (dashed line) giving the electric field. **c**, The spin Hall effect. The excess of spin-up electrons (red arrows) at the cool end of the ferromagnet drives a spin current that flows vertically into the platinum film (yellow arrow). Here, spin-up means that the direction of the spin is parallel to the magnetic field, and thus points to the right. By spin-orbit coupling, electrons 'see' a weak magnetic field circulating around a charged impurity (circles around blue dot). Scattering from the charged impurity causes spin-up electrons to accumulate preferentially on the far face of the platinum film, whereas spin-down electrons (blue arrows) end up on the near face. The imbalance is observed as a Hall voltage difference between the two faces.

drives a spin current into the platinum film. The decay of this spin current leads to the appearance of an electrical signal through the spin Hall effect^{3,4}, which enables spin currents to be detected using a sensitive voltmeter.

To understand the spin Hall effect, one can track an electron as it enters the platinum film and scatters off a charged impurity (Fig. 1c). The spin-orbit interaction — the interaction between an electron's spin and its motion — imparts a left-right asymmetry to the scattering

process. In the rest frame of the electron, the charged impurity rushing towards it constitutes a current filament, so the electron 'sees' a weak magnetic field circling the filament. This non-uniform magnetic field imparts a force on the electron along a direction that depends on its spin orientation^{3,4}. The net result is that spin-up electrons are pushed to the right of the impurity whereas spin-down electrons are pushed to its left.

In effect, each impurity acts like a spin filter that selectively kicks electrons to one side or the other, depending on their spin. As shown in Figure 1c, the excess spin-up population in the incident beam results in more charge accumulating on the far face than on the near face of the platinum film. The voltage difference between the two faces is observable as a Hall signal. The asymmetric scattering of electrons is especially large in materials with a high atomic number, such as platinum. Following its prediction^{3,4}, the spin Hall effect was first observed by applying purely optical techniques to semiconductors^{5,6}, and was later detected electrically in metals^{7,8}.

In a series of tests, Uchida *et al.*² convincingly show that the Hall voltage in the platinum film arises from the spin voltage. The Hall signal in

the platinum film tracks both the magnitude and the direction of the magnetization in the nickel-iron film. Moreover, by moving the platinum film along the length of the nickel-iron film, they show that the spin voltage varies linearly over the 6-mm length of the sample. In demonstrating that the spin Seebeck effect can produce a large, calibrated spin-voltage source that can be 'tapped' anywhere along the length of the ferromagnet, Uchida and colleagues have added an important tool to the spintronics toolbox. ■

N. P. Ong is in the Department of Physics, Princeton University, Princeton, New Jersey 08544, USA.

e-mail: npo@princeton.edu

1. Gregg, J. F. in *Spin Electronics* (eds Ziese, M. & Thornton, M. J.) 3–31 (Springer, 2001).
2. Uchida, K. *et al.* *Nature* **455**, 778–781 (2008).
3. D'yakonov, M. I. & Perel, V. I. *Phys. Lett. A* **35**, 459–460 (1971).
4. Hirsch, J. E. *Phys. Rev. Lett.* **83**, 1834–1837 (1999).
5. Kato, Y. K., Myers, R. C., Gossard, A. C. & Awschalom, D. D. *Science* **306**, 1910–1913 (2004).
6. Wunderlich, J., Kaestner, B., Sinova, J. & Jungwirth, T. *Phys. Rev. Lett.* **94**, 047204 (2005).
7. Valenzuela, S. O. & Tinkham, M. *Nature* **442**, 176–179 (2006).
8. Kimura, T., Otani, Y., Sato, T., Takahashi, S. & Maekawa, S. *Phys. Rev. Lett.* **98**, 156601 (2007).

NEUROSCIENCE

Brain's defence against cocaine

L. Judson Chandler and Peter W. Kalivas

Long-term exposure to cocaine changes the organization of synaptic connections within the addiction circuitry of the brain. This process might protect against the development and persistence of addiction.

Neurons modify their structure and communication with other neurons in response to experiences. Such experience-dependent neuroplasticity is crucial for survival because it allows learning from, and responses to, changes in the environment. But the cellular mechanisms that mediate this process can also be co-opted by drugs of abuse. Reporting in *Neuron*, Pulipparacharuvil *et al.*¹ describe how some of the chemical, structural and behavioural changes in neurons that are induced by repeated exposure to cocaine are regulated at a molecular level.

Drug addiction is characterized by compulsive drug seeking. It resembles a chronic relapsing disorder in which the addict resumes taking drugs after a period or periods of abstinence. Human and animal studies indicate that the recalcitrant nature of addiction results from drug-induced stimulation of reward-related learning processes in the brain. The pleasure-producing effects of the drug trigger cellular and molecular processes that are normally activated by natural rewards such as food and sex. Repeated exposure to an addictive drug

leads to a long-lasting associative memory of its rewarding properties through experience-dependent neuroplasticity. In effect, drug-seeking behaviour becomes hard-wired in the addict's brain, and the persistent memory trace is easily reactivated by drug-associated environmental stimuli, such as the sight of drug paraphernalia.

Along the dendritic processes of a neuron, morphologically specialized structures called dendritic spines receive most of the excitatory signals from other neurons through synaptic junctions. These spines are considered to be a primary cellular site for mediating the synaptic plasticity that is thought to underpin memory formation². One regulator of the density of excitatory signals on dendritic spines is the gene transcription factor MEF2 (ref. 3). When active (dephosphorylated), MEF2 favours elimination of dendritic spines, and when inactive (phosphorylated) it allows spine formation^{3,4}.

Repeated exposure to cocaine and other psychostimulants increases the number of dendritic spines on medium spiny neurons

MATERIALS SCIENCE

Solar cells go round the bend

With high oil prices sparking a surge of interest in alternative energy sources, solar cells have become the subject of intense research. Much of this effort focuses on finding new designs that open up fresh applications. John Rogers and colleagues now report just such a development (J. Yoon *et al.* *Nature Mater.* doi:10.1038/nmat2287; 2008) — tiny, ultrathin cells made of silicon that, when fixed in arrays on a flexible substrate, create large, bendy solar cells (pictured).

The authors carve their microcell arrays from a rectangular block of silicon. They begin by etching the outlines of the microcells (the tops and sides) onto the upper surface of the silicon block. They then make electronic junctions and electrical contacts by 'doping' the silicon,

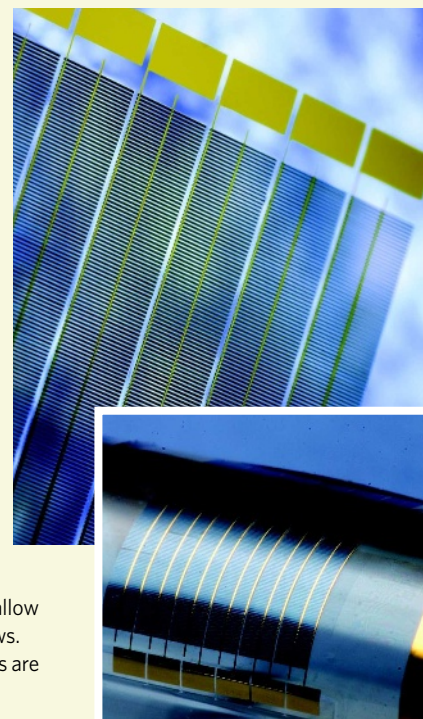
adding boron and phosphorus, and using an inert mask to define the regions to be doped. A further round of etching exposes the final three-dimensional shape of the microcells, retaining a thin sliver of silicon to anchor the cells to the block. Finally, the base of the wafer is doped with boron, to yield functioning solar microcells.

To make bendable, large-scale solar cells, Rogers and colleagues use a printing technique. They press a flat stamp onto the arrays of microcells on the silicon block, breaking the anchors that tether them to the silicon. The microcells stick to the soft surface of the stamp, and are transferred to a flexible substrate simply by pressing the stamp onto the substrate. The authors then construct electrodes to

connect the microcells to each other, using one of various established methods.

The resulting devices have several desirable properties. First, they are remarkably light, which, along with their flexibility, allows them to be transported and installed more easily than existing solar cells. Second, they work just as efficiently when bent as they do when flat, so they could be fixed to curved or irregular surfaces. Furthermore, they can be made to be transparent, which would allow them to be used on windows. And because the microcells are so thin, less silicon is used, minimizing costs.

Andrew Mitchinson



J. YOON ET AL.

in the nucleus accumbens — a primary brain region in the addiction neurocircuitry. So Pulipparacharuvil and colleagues¹ hypothesized that such cocaine-induced structural neuroplasticity might also be regulated by MEF2. Indeed, they demonstrate that this transcription factor, which is highly expressed in medium spiny neurons and is predominantly active under normal conditions¹, is affected by chronic cocaine administration. Specifically, long-term exposure to cocaine seems to prevent MEF2 dephosphorylation by a calcium/calmodulin-dependent phosphatase enzyme known as calcineurin, thereby suppressing its activation.

Cocaine-induced MEF2 inhibition also seems to involve enhanced phosphorylation of this transcription factor by a kinase known as Cdk5. In the nucleus accumbens, Cdk5 activity — which modulates behavioural responses to cocaine, such as motivation to consume the drug⁵ — increases after chronic exposure to cocaine⁶. Together with Pulipparacharuvil and colleagues' data, these observations^{4–6} strongly suggest that repeated exposure to cocaine inhibits MEF2 activity through both enhanced phosphorylation by Cdk5 and attenuation of dephosphorylation by calcineurin. The reduction in MEF2's transcriptional activity in turn promotes increases in the number of dendritic spines.

The basic mechanism underlying experience-dependent synaptic plasticity is often described by the phrase "Neurons that fire together, wire together"⁷. Reward-related associative learning is a form of such 'Hebbian plasticity', in which synaptic connections are enhanced by the improved strength of existing

synapses and/or by an increase in the number of such connections. So a logical conclusion would be that cocaine-induced increases in spine density reflect an activity-dependent strengthening of synaptic connectivity, which presumably underlies addictive behaviour. Surprisingly, however, Pulipparacharuvil and colleagues' observations¹ do not support this inference. By manipulating MEF2 activity, they inhibited cocaine-induced increases in spine density. However, this did not seem to prevent increases in the behavioural response to this drug, and might even promote it. So increases in spine density resulting from MEF2 inhibition seem to be associated with reduced behavioural sensitivity to cocaine.

If bulk increase in spine density within the nucleus accumbens does not contribute to enhanced behavioural responses to cocaine, then what is its function, and how can it be reconciled with the processes of experience-dependent associative learning? One confounding aspect of Hebbian plasticity is that, when allowed to proceed unchecked, activity-dependent changes in synaptic connections can destabilize neural networks⁸. In self-defence, the brain uses homeostatic-plasticity mechanisms to oppose such destabilizing effects.

Homeostatic plasticity tends to occur on a large scale to maintain the overall firing activity of a neuron. This allows synapse-specific remodelling of neuronal circuits to proceed through Hebbian mechanisms while maintaining stability of the overall neural network. A major excitatory input reaching medium spiny neurons originates from the prefrontal cortex,

but chronic exposure to cocaine markedly alters the output of prefrontal cortical neurons projecting to the nucleus accumbens⁹. So, as Pulipparacharuvil *et al.* suggest, an intriguing possibility is that cocaine-induced increases in spine density in the nucleus accumbens, which are mediated by MEF2 inhibition, may represent a homeostatic response to altered excitatory input from the prefrontal cortex.

These findings¹ provide a direct challenge to the view that increased spine density induced by repeated exposure to psychostimulants underlies maladaptive plasticity. Moreover, they agree with previous observations¹⁰ that identified compensatory drug-induced neuroadaptations. Future research into the neuroplasticity induced by addictive drugs must therefore consider competition between activity-dependent remodelling of synaptic connections and homeostatic adaptations that maintain overall stability in neuronal networks. ■

L. Judson Chandler and Peter W. Kalivas are in the Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina 29425, USA.
e-mails: chandj@musc.edu; kalivasp@musc.edu

1. Pulipparacharuvil, S. *et al.* *Neuron* **59**, 621–633 (2008).
2. Alvarez, V. A. & Sabatini, B. L. *Annu. Rev. Neurosci.* **30**, 79–97 (2007).
3. Flavell, S. W. *et al.* *Science* **311**, 1008–1012 (2006).
4. Gong, X. *et al.* *Neuron* **38**, 33–46 (2003).
5. Benavides, D. R. *et al.* *J. Neurosci.* **27**, 12967–12976 (2007).
6. Bibb, J. A. *et al.* *Nature* **410**, 376–380 (2001).
7. Bi, G. & Poo, M. *Annu. Rev. Neurosci.* **24**, 139–166 (2001).
8. Abbott, L. F. & Nelson, S. B. *Nature Neurosci.* **3**, 1178–1183 (2000).
9. Kalivas, P. W., Volkow, N. & Seamans, J. *Neuron* **45**, 647–650 (2005).
10. Toda, S. *et al.* *J. Neurosci.* **26**, 1579–1587 (2006).

MATERIALS SCIENCE

Solar cells go round the bend

With high oil prices sparking a surge of interest in alternative energy sources, solar cells have become the subject of intense research. Much of this effort focuses on finding new designs that open up fresh applications. John Rogers and colleagues now report just such a development (J. Yoon *et al. Nature Mater.* doi:10.1038/nmat2287; 2008) — tiny, ultrathin cells made of silicon that, when fixed in arrays on a flexible substrate, create large, bendy solar cells (pictured).

The authors carve their microcell arrays from a rectangular block of silicon. They begin by etching the outlines of the microcells (the tops and sides) onto the upper surface of the silicon block. They then make electronic junctions and electrical contacts by 'doping' the silicon,

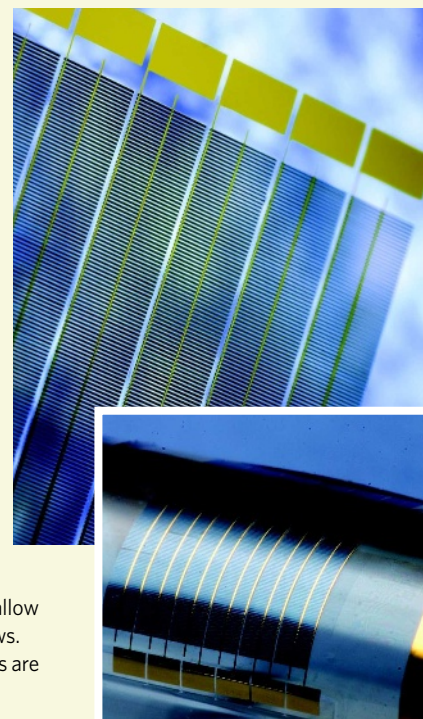
adding boron and phosphorus, and using an inert mask to define the regions to be doped. A further round of etching exposes the final three-dimensional shape of the microcells, retaining a thin sliver of silicon to anchor the cells to the block. Finally, the base of the wafer is doped with boron, to yield functioning solar microcells.

To make bendable, large-scale solar cells, Rogers and colleagues use a printing technique. They press a flat stamp onto the arrays of microcells on the silicon block, breaking the anchors that tether them to the silicon. The microcells stick to the soft surface of the stamp, and are transferred to a flexible substrate simply by pressing the stamp onto the substrate. The authors then construct electrodes to

connect the microcells to each other, using one of various established methods.

The resulting devices have several desirable properties. First, they are remarkably light, which, along with their flexibility, allows them to be transported and installed more easily than existing solar cells. Second, they work just as efficiently when bent as they do when flat, so they could be fixed to curved or irregular surfaces. Furthermore, they can be made to be transparent, which would allow them to be used on windows. And because the microcells are so thin, less silicon is used, minimizing costs.

Andrew Mitchinson



J. YOON ET AL.

in the nucleus accumbens — a primary brain region in the addiction neurocircuitry. So Pulipparacharuvil and colleagues¹ hypothesized that such cocaine-induced structural neuroplasticity might also be regulated by MEF2. Indeed, they demonstrate that this transcription factor, which is highly expressed in medium spiny neurons and is predominantly active under normal conditions¹, is affected by chronic cocaine administration. Specifically, long-term exposure to cocaine seems to prevent MEF2 dephosphorylation by a calcium/calmodulin-dependent phosphatase enzyme known as calcineurin, thereby suppressing its activation.

Cocaine-induced MEF2 inhibition also seems to involve enhanced phosphorylation of this transcription factor by a kinase known as Cdk5. In the nucleus accumbens, Cdk5 activity — which modulates behavioural responses to cocaine, such as motivation to consume the drug⁵ — increases after chronic exposure to cocaine⁶. Together with Pulipparacharuvil and colleagues' data, these observations^{4–6} strongly suggest that repeated exposure to cocaine inhibits MEF2 activity through both enhanced phosphorylation by Cdk5 and attenuation of dephosphorylation by calcineurin. The reduction in MEF2's transcriptional activity in turn promotes increases in the number of dendritic spines.

The basic mechanism underlying experience-dependent synaptic plasticity is often described by the phrase "Neurons that fire together, wire together"⁷. Reward-related associative learning is a form of such 'Hebbian plasticity', in which synaptic connections are enhanced by the improved strength of existing

synapses and/or by an increase in the number of such connections. So a logical conclusion would be that cocaine-induced increases in spine density reflect an activity-dependent strengthening of synaptic connectivity, which presumably underlies addictive behaviour. Surprisingly, however, Pulipparacharuvil and colleagues' observations¹ do not support this inference. By manipulating MEF2 activity, they inhibited cocaine-induced increases in spine density. However, this did not seem to prevent increases in the behavioural response to this drug, and might even promote it. So increases in spine density resulting from MEF2 inhibition seem to be associated with reduced behavioural sensitivity to cocaine.

If bulk increase in spine density within the nucleus accumbens does not contribute to enhanced behavioural responses to cocaine, then what is its function, and how can it be reconciled with the processes of experience-dependent associative learning? One confounding aspect of Hebbian plasticity is that, when allowed to proceed unchecked, activity-dependent changes in synaptic connections can destabilize neural networks⁸. In self-defence, the brain uses homeostatic-plasticity mechanisms to oppose such destabilizing effects.

Homeostatic plasticity tends to occur on a large scale to maintain the overall firing activity of a neuron. This allows synapse-specific remodelling of neuronal circuits to proceed through Hebbian mechanisms while maintaining stability of the overall neural network. A major excitatory input reaching medium spiny neurons originates from the prefrontal cortex,

but chronic exposure to cocaine markedly alters the output of prefrontal cortical neurons projecting to the nucleus accumbens⁹. So, as Pulipparacharuvil *et al.* suggest, an intriguing possibility is that cocaine-induced increases in spine density in the nucleus accumbens, which are mediated by MEF2 inhibition, may represent a homeostatic response to altered excitatory input from the prefrontal cortex.

These findings¹ provide a direct challenge to the view that increased spine density induced by repeated exposure to psychostimulants underlies maladaptive plasticity. Moreover, they agree with previous observations¹⁰ that identified compensatory drug-induced neuroadaptations. Future research into the neuroplasticity induced by addictive drugs must therefore consider competition between activity-dependent remodelling of synaptic connections and homeostatic adaptations that maintain overall stability in neuronal networks. ■

L. Judson Chandler and Peter W. Kalivas are in the Department of Neurosciences, Medical University of South Carolina, Charleston, South Carolina 29425, USA.
e-mails: chandj@musc.edu; kalivasp@musc.edu

1. Pulipparacharuvil, S. *et al. Neuron* **59**, 621–633 (2008).
2. Alvarez, V. A. & Sabatini, B. L. *Annu. Rev. Neurosci.* **30**, 79–97 (2007).
3. Flavell, S. W. *et al. Science* **311**, 1008–1012 (2006).
4. Gong, X. *et al. Neuron* **38**, 33–46 (2003).
5. Benavides, D. R. *et al. J. Neurosci.* **27**, 12967–12976 (2007).
6. Bibb, J. A. *et al. Nature* **410**, 376–380 (2001).
7. Bi, G. & Poo, M. *Annu. Rev. Neurosci.* **24**, 139–166 (2001).
8. Abbott, L. F. & Nelson, S. B. *Nature Neurosci.* **3**, 1178–1183 (2000).
9. Kalivas, P. W., Volkow, N. & Seamans, J. *Neuron* **45**, 647–650 (2005).
10. Toda, S. *et al. J. Neurosci.* **26**, 1579–1587 (2006).

IMMUNOLOGY

Helpful T cells are sticky

Elissa K. Deenick and Stuart G. Tangye

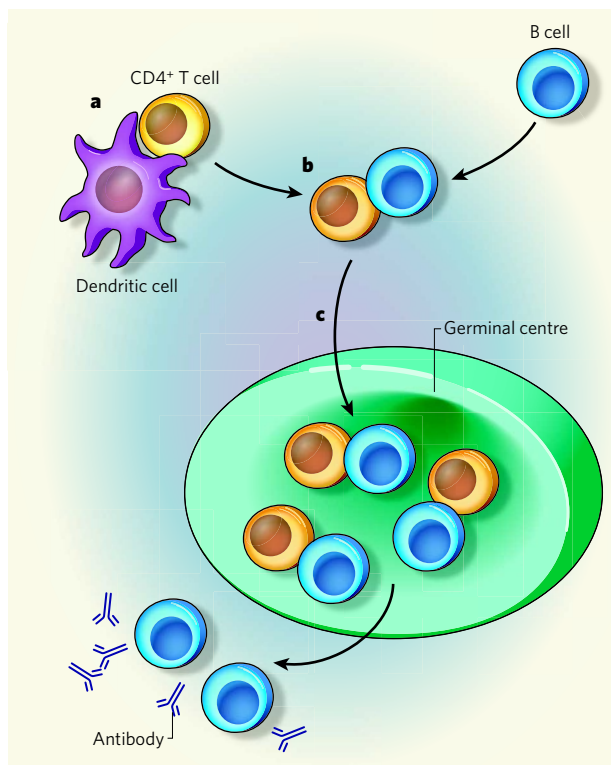
Prolonged physical interaction between helper T cells and antibody-producing B cells is crucial for efficient immune responses. Mutations in a protein that underlies this process cause human disease.

The production of antibodies by B cells is essential for protective immunity following vaccination or exposure to infectious pathogens. The development of antibody-secreting B cells occurs in discrete areas of lymphoid tissues called germinal centres^{1,2}, the formation of which depends on interactions between B cells and T cells bearing the CD4 molecule on their surface (CD4⁺ T cells). But several steps in the orchestration of T- and B-cell activation, differentiation and 'homing' to germinal centres during an immune response remain incompletely defined. For example, mutations in the protein SAP, which is involved in signalling by the SLAM family of cell-surface receptors³, leads to defects in the formation of germinal centres and the generation of long-lived antibody-secreting B cells. These defects result in a human immunodeficiency condition called X-linked lymphoproliferative disease³. But the mechanism associated with loss of SAP function has remained unknown. On page 764 of this issue, Qi *et al.*⁴ shed light on how SAP functions in CD4⁺ T cells to efficiently engage B cells and to provide appropriate signals for both the formation of germinal centres and the differentiation of B cells into antibody-secreting cells.

Several specialized immune cells mediate B-cell differentiation into long-lived antibody-secreting cells. Initially, dendritic cells capture foreign antigens and present them to CD4⁺ T cells, thereby activating them^{1,2}. These antigen-specific CD4⁺ T helper cells then interact with antigen-specific B cells, which undergo intense proliferation and eventually differentiate into long-lived antibody-secreting cells (Fig. 1). Qi *et al.* show that SAP is not required for dendritic cells to bind to and activate CD4⁺ T cells. Instead, the authors find that SAP-deficient, activated T cells cannot form stable interactions with B cells. The reduction in contact time between T and B cells probably explains the failure of SAP-deficient CD4⁺ T cells to deliver the necessary contact-mediated helper signals to B cells.

The authors also find that activated SAP-deficient CD4⁺ T cells show characteristics of functional follicular helper T (T_{FH}) cells, which are normally found in the germinal centres⁵. For instance, like T_{FH} cells, CD4⁺ T cells from SAP-deficient mice express high levels of specific surface molecules, including CXCR5, ICOS, CD40L and OX40. Nonetheless, these cells fail to efficiently enter or remain within germinal centres — a central requirement for

Figure 1 | Role of SAP in immune responses mediated by T and B cells. **a**, In lymphoid tissues, antigens on dendritic cells lead to these cells binding to, and activating, CD4⁺ T helper cells. **b**, The activated T cells then interact with antigen-specific B cells. Qi *et al.*⁴ show that, in the absence of the SAP protein, T and B cells manage only brief interactions. **c**, Activated B cells then form germinal centres, where they continue to receive help from CD4⁺ T cells. SAP-deficient T cells have limited ability to enter and remain in the germinal centres.



50 YEARS AGO

So far as men of science are concerned, the Lambeth Conference report follows much the same pattern as its immediate predecessor. The bulk of the report is concerned with topics which are not the immediate concern of scientists, as such, though they will note the resolution which gratefully acknowledges the work of scientists in increasing man's knowledge of the universe ... [T]wo sections of the report are concerned with problems with which men of science, as such, are equally concerned ... First are the problems involved in reconciliation of the conflicts between and within nations, and second are the group of problems centring around the family in contemporary society ... [T]he concluding section of the report, in which political conflicts are considered, merits close attention, because it poses a problem of action and of impartiality with which scientists are themselves familiar and which lies at the root of any attempt to apply scientific or technological knowledge impartially and objectively in public affairs.

From *Nature* 11 October 1958.

100 YEARS AGO

(1) *Selectionsprinzip und Probleme der Artbildung*: ein Handbuch des Darwinismus. By Prof. Ludwig Plate; (2) *Die Lehre Darwins in ihren letzten Folgen*. By Max Steiner — Prof. L. Plate's "Selectionsprinzip" has been so much expanded in its third edition that it deserves to be called a "handbook of Darwinism". It is a careful and thoughtful text-book by a thorough-going Darwinian, who is at the same time a believer in the transmission of acquired characters ... The author of the second volume before us seems to think that Darwinism has been too much discussed as a biological theory, artificially abstracted from its social consequences. If we understand him, he seeks to put things right by showing what terrible consequences the theory involves.

From *Nature* 15 October 1908.

50 & 100 YEARS AGO

a T_{HH} cell to fulfil its duty of helping B cells^{5,6}. So, in X-linked lymphoproliferative disease, defects in germinal-centre formation and antibody production seem to be due not only to inadequate communication between T and B cells but also to failed homing of T_{HH} cells to the germinal centres.

These findings have two noteworthy implications. First, they indicate that CD4⁺ T cells use different sets of molecules for each of the cell types with which they communicate and interact. Specifically, SAP — and, by inference, the SLAM family of cell-surface receptors — is required for the dialogue between CD4⁺ T cells and B cells but not for that between T cells and dendritic cells. Indeed, increased expression of specific SLAM proteins (CD84, SLAM, Ly108 and CD229) on B cells but not on dendritic cells⁴ supports this conclusion.

Second, the data⁴ suggest that the array of molecules involved in the dialogue between dendritic cells and T cells is insufficient to induce functional T_{HH} cells. Instead, it seems that B cells provide a unique signal that allows the appropriate CD4⁺ T cells to become fully functional T_{HH} cells — an idea supported by work in B-cell-deficient mice⁷. By inference, therefore, the definition of T_{HH} cells should be refined beyond their expression of molecules such as CXCR5. Indeed, earlier studies^{6,8} noted that the population of CXCR5-expressing cells includes CD4⁺ T cells found not only in germinal centres, but also outside them. Future work should determine the contributions of these different CXCR5-expressing CD4⁺ T-cell populations to B-cell responses and identify more specifically the T_{HH} cells that are truly located in germinal centres.

SAP binds to the cytoplasmic domain of SLAM-family cell-surface receptors. A crucial question arising from Qi and colleagues' study⁴ is which SLAM members are required for optimal adhesion of T cells to B cells. Although SLAM and CD229 are highly expressed on B cells, their deletion does not impair germinal-centre formation or T-cell-dependent antibody responses^{9,10}. CD84, however, could be a promising candidate, as it is highly expressed on both T_{HH} and B cells^{3–5,11}. So (presumably SAP-dependent) interactions between CD84 molecules on these cells might contribute to the formation of stable conjugates between T_{HH} and germinal-centre B cells, which seem to be essential for the efficient production of antibodies. Generation of CD84-deficient mice will clarify the role of this receptor in mediating interactions between T and B cells.

How does SAP itself contribute to adhesion between T and B cells? SAP-dependent signalling downstream of the SLAM-family receptors may induce changes in the expression of other adhesion molecules, such as integrins, that are involved in interactions between T and B cells. But the introduction of a signalling-deficient version of SAP into SAP-deficient CD4⁺ T cells can restore adhesion between B and T cells⁴ — an observation that hints that signalling

through SAP-associating receptors per se is not required for normal interactions between these cells. Alternatively, SLAM-family members may operate as adhesion molecules only in the presence of functional SAP (ref. 3). In other words, although SAP is unlikely to regulate the expression levels of SLAM receptors, it might stabilize interactions between these receptors on B cells and CD4⁺ T cells.

In mice, genes encoding SLAM-family receptors lie in a region known to be associated with susceptibility to the autoimmune disease systemic lupus erythematosus¹². So Qi and colleagues' results also have potential implications for understanding autoimmune diseases. Variations in the genes encoding SLAM proteins are predicted¹² to influence the strength of interactions between the extracellular domains of these cell-surface receptors or between their cytoplasmic domains and SAP. If reduced adhesion between B cells and SAP-deficient T_{HH} cells contributes to immunodeficiency, as occurs in X-linked lymphoproliferative disease, the converse — prolonged interactions between T and B cells through increased binding strength — might result in amplified T-cell

help and abnormal antibody responses characteristic of autoimmunity. By revealing more of the steps in the intricate dance of collaboration between T and B cells leading to antibody production, this study⁴ provides potential routes for modulating aberrant immunity in both immunodeficiency and autoimmunity. ■

Elissa K. Deenick and Stuart G. Tangye are at the Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, 2010 New South Wales, Australia. e-mails: e.deenick@garvan.org.au; s.tangye@garvan.org.au

1. Liu, Y. J. & Banchereau, J. *Immunologist* **4**, 55–66 (1996).
2. MacLennan, I. C. M. *Annu. Rev. Immunol.* **12**, 117–139 (1994).
3. Ma, C. S., Nichols, K. E. & Tangye, S. G. *Annu. Rev. Immunol.* **25**, 337–379 (2007).
4. Qi, H., Cannons, J. L., Klauschen, F., Schwartzberg, P. L. & Germain, R. N. *Nature* **455**, 764–769 (2008).
5. Vinuesa, C. G., Tangye, S. G., Moser, B. & Mackay, C. R. *Nature Rev. Immunol.* **5**, 853–865 (2005).
6. Kim, C. H. et al. *J. Exp. Med.* **193**, 1373–1381 (2001).
7. Haynes, N. M. et al. *J. Immunol.* **179**, 5099–5108 (2007).
8. Ansel, K. M. et al. *J. Exp. Med.* **190**, 1123–1134 (1999).
9. Graham, D. B. et al. *J. Immunol.* **176**, 291–300 (2006).
10. McCausland, M. M. et al. *J. Immunol.* **178**, 817–828 (2007).
11. Chtanova, T. et al. *J. Immunol.* **173**, 68–78 (2004).
12. Chan, A. Y. et al. *Curr. Opin. Immunol.* **18**, 656–664 (2006).

DEVELOPMENTAL BIOLOGY

Teeth in double trouble

Georgy Koentges

Almost all vertebrates have teeth of some sort. But where, in developmental terms, do teeth come from? Results drawn from experimental embryology provide an illuminating perspective on this contentious question.

Teeth are made of some of the hardest stuff in organic nature, and many fossil vertebrates are known only from their dental remains. So teeth are central for systematic classification and reconstruction of animal life-histories, not to mention forensic science, horror movies and musicals. But we know all too little about the earliest cellular and molecular events that initiate teeth and define their position, shape and patterns — a deficiency that Soukup *et al.* (page 795 of this issue¹) have set out to remedy by first sorting out some basic embryology.

Three cell lineages in the vertebrate embryo pertain to tooth development — ectoderm and endoderm, organized as epithelia, and mesenchyme, derived from the so-called neural crest. Tissue interactions between embryonic epithelia and mesenchyme are known to be needed to form teeth². In all bony fish, for example, the epithelia form specialized cells that make the tooth enamel, whereas the mesenchyme makes the underlying dentine. But vertebrate hard tissues are complex: the same neural-crest cells can also form bone, and it is not known how such differences are established. A substantial body of work³ has elucidated the molecular details of downstream signalling

systems that sculpt teeth. But the very earliest events that determine tooth patterning remain obscure.

In evolutionary terms, tooth-like structures — such as the denticles that appear as a ubiquitous feature on the body armour of early vertebrates — might have preceded the advent of jaws proper⁴. The staggering histological diversity of such structures has led to byzantine systems of classification of vertebrate hard tissues, and in turn to serious differences of opinion. The acrimony of these debates has scaled linearly with the lack of experimental embryological evidence about the underlying process.

The presence of denticles on the body of early jawed vertebrates led to speculation that, early in vertebrate evolution, embryonic ectoderm moved into the mouth and initiated organized tooth rows there. In contrast to this 'outside-in' view of events is the 'inside-out' theory. This theory holds that the evolutionary origins of teeth started in the mouth or pharynx and are linked to the presence of embryonic endoderm. An outward migration of cells, or a co-option of a pharyngeal tooth-forming program in a part of the outer body surface,

a T_{FH} cell to fulfil its duty of helping B cells^{5,6}. So, in X-linked lymphoproliferative disease, defects in germinal-centre formation and antibody production seem to be due not only to inadequate communication between T and B cells but also to failed homing of T_{FH} cells to the germinal centres.

These findings have two noteworthy implications. First, they indicate that $CD4^+$ T cells use different sets of molecules for each of the cell types with which they communicate and interact. Specifically, SAP — and, by inference, the SLAM family of cell-surface receptors — is required for the dialogue between $CD4^+$ T cells and B cells but not for that between T cells and dendritic cells. Indeed, increased expression of specific SLAM proteins (CD84, SLAM, Ly108 and CD229) on B cells but not on dendritic cells⁴ supports this conclusion.

Second, the data⁴ suggest that the array of molecules involved in the dialogue between dendritic cells and T cells is insufficient to induce functional T_{FH} cells. Instead, it seems that B cells provide a unique signal that allows the appropriate $CD4^+$ T cells to become fully functional T_{FH} cells — an idea supported by work in B-cell-deficient mice⁷. By inference, therefore, the definition of T_{FH} cells should be refined beyond their expression of molecules such as CXCR5. Indeed, earlier studies^{6,8} noted that the population of CXCR5-expressing cells includes $CD4^+$ T cells found not only in germinal centres, but also outside them. Future work should determine the contributions of these different CXCR5-expressing $CD4^+$ T-cell populations to B-cell responses and identify more specifically the T_{FH} cells that are truly located in germinal centres.

SAP binds to the cytoplasmic domain of SLAM-family cell-surface receptors. A crucial question arising from Qi and colleagues' study⁴ is which SLAM members are required for optimal adhesion of T cells to B cells. Although SLAM and CD229 are highly expressed on B cells, their deletion does not impair germinal-centre formation or T-cell-dependent antibody responses^{9,10}. CD84, however, could be a promising candidate, as it is highly expressed on both T_{FH} and B cells^{3-5,11}. So (presumably SAP-dependent) interactions between CD84 molecules on these cells might contribute to the formation of stable conjugates between T_{FH} and germinal-centre B cells, which seem to be essential for the efficient production of antibodies. Generation of CD84-deficient mice will clarify the role of this receptor in mediating interactions between T and B cells.

How does SAP itself contribute to adhesion between T and B cells? SAP-dependent signalling downstream of the SLAM-family receptors may induce changes in the expression of other adhesion molecules, such as integrins, that are involved in interactions between T and B cells. But the introduction of a signalling-deficient version of SAP into SAP-deficient $CD4^+$ T cells can restore adhesion between B and T cells⁴ — an observation that hints that signalling

through SAP-associating receptors per se is not required for normal interactions between these cells. Alternatively, SLAM-family members may operate as adhesion molecules only in the presence of functional SAP (ref. 3). In other words, although SAP is unlikely to regulate the expression levels of SLAM receptors, it might stabilize interactions between these receptors on B cells and $CD4^+$ T cells.

In mice, genes encoding SLAM-family receptors lie in a region known to be associated with susceptibility to the autoimmune disease systemic lupus erythematosus¹². So Qi and colleagues' results also have potential implications for understanding autoimmune diseases. Variations in the genes encoding SLAM proteins are predicted¹² to influence the strength of interactions between the extracellular domains of these cell-surface receptors or between their cytoplasmic domains and SAP. If reduced adhesion between B cells and SAP-deficient T_{FH} cells contributes to immunodeficiency, as occurs in X-linked lymphoproliferative disease, the converse — prolonged interactions between T and B cells through increased binding strength — might result in amplified T-cell

help and abnormal antibody responses characteristic of autoimmunity. By revealing more of the steps in the intricate dance of collaboration between T and B cells leading to antibody production, this study⁴ provides potential routes for modulating aberrant immunity in both immunodeficiency and autoimmunity. ■

Elissa K. Deenick and Stuart G. Tangye are at the Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, 2010 New South Wales, Australia. e-mails: e.deenick@garvan.org.au; s.tangye@garvan.org.au

1. Liu, Y. J. & Banchereau, J. *Immunologist* **4**, 55–66 (1996).
2. MacLennan, I. C. M. *Annu. Rev. Immunol.* **12**, 117–139 (1994).
3. Ma, C. S., Nichols, K. E. & Tangye, S. G. *Annu. Rev. Immunol.* **25**, 337–379 (2007).
4. Qi, H., Cannons, J. L., Klauschen, F., Schwartzberg, P. L. & Germain, R. N. *Nature* **455**, 764–769 (2008).
5. Vinuesa, C. G., Tangye, S. G., Moser, B. & Mackay, C. R. *Nature Rev. Immunol.* **5**, 853–865 (2005).
6. Kim, C. H. et al. *J. Exp. Med.* **193**, 1373–1381 (2001).
7. Haynes, N. M. et al. *J. Immunol.* **179**, 5099–5108 (2007).
8. Ansel, K. M. et al. *J. Exp. Med.* **190**, 1123–1134 (1999).
9. Graham, D. B. et al. *J. Immunol.* **176**, 291–300 (2006).
10. McCausland, M. M. et al. *J. Immunol.* **178**, 817–828 (2007).
11. Chtanova, T. et al. *J. Immunol.* **173**, 68–78 (2004).
12. Chan, A. Y. et al. *Curr. Opin. Immunol.* **18**, 656–664 (2006).

DEVELOPMENTAL BIOLOGY

Teeth in double trouble

Georgy Koentges

Almost all vertebrates have teeth of some sort. But where, in developmental terms, do teeth come from? Results drawn from experimental embryology provide an illuminating perspective on this contentious question.

Teeth are made of some of the hardest stuff in organic nature, and many fossil vertebrates are known only from their dental remains. So teeth are central for systematic classification and reconstruction of animal life-histories, not to mention forensic science, horror movies and musicals. But we know all too little about the earliest cellular and molecular events that initiate teeth and define their position, shape and patterns — a deficiency that Soukup *et al.* (page 795 of this issue¹) have set out to remedy by first sorting out some basic embryology.

Three cell lineages in the vertebrate embryo pertain to tooth development — ectoderm and endoderm, organized as epithelia, and mesenchyme, derived from the so-called neural crest. Tissue interactions between embryonic epithelia and mesenchyme are known to be needed to form teeth². In all bony fish, for example, the epithelia form specialized cells that make the tooth enamel, whereas the mesenchyme makes the underlying dentine. But vertebrate hard tissues are complex: the same neural-crest cells can also form bone, and it is not known how such differences are established. A substantial body of work³ has elucidated the molecular details of downstream signalling

systems that sculpt teeth. But the very earliest events that determine tooth patterning remain obscure.

In evolutionary terms, tooth-like structures — such as the denticles that appear as a ubiquitous feature on the body armour of early vertebrates — might have preceded the advent of jaws proper⁴. The staggering histological diversity of such structures has led to byzantine systems of classification of vertebrate hard tissues, and in turn to serious differences of opinion. The acrimony of these debates has scaled linearly with the lack of experimental embryological evidence about the underlying process.

The presence of denticles on the body of early jawed vertebrates led to speculation that, early in vertebrate evolution, embryonic ectoderm moved into the mouth and initiated organized tooth rows there. In contrast to this 'outside-in' view of events is the 'inside-out' theory. This theory holds that the evolutionary origins of teeth started in the mouth or pharynx and are linked to the presence of embryonic endoderm. An outward migration of cells, or a co-option of a pharyngeal tooth-forming program in a part of the outer body surface,

would have to occur to explain the presence of denticles on the outer covering of sharks and other more basal vertebrates^{5,6}.

Both theories hinge on the idea that there is an inherent difference in the inductive power of ectoderm and endoderm, and that migration of one or the other is the crucial factor in tooth formation. Implicit in this is the notion that tooth and denticle anatomy reflects embryonic origins — that is, that actual tooth or denticle histology can reveal which embryonic tissue was the key source.

Soukup *et al.*¹ now provide experimental grounds to debunk such ideas by testing the spatial distribution of ectoderm and endoderm in relation to erupting teeth. They took advantage of a line of transgenic axolotls⁷ — a group of amphibians that have ample teeth in their mouth — in which cell lineages can be fluorescently labelled. By grafting ectoderm or endoderm labelled with a marker known as green fluorescent protein from transgenic into normal axolotls, and vice versa, the authors show that there is no relationship between ectodermal and endodermal origin and the shape or nature of the resulting teeth — at least at the point when such teeth become visible. The enamel of teeth can be of ectodermal, endodermal or mixed origin. This is a dramatic finding. It means that one cannot infer relative distributions of ectoderm and endoderm from tooth or denticle anatomy even in a living species, let alone in a fossil.

The caveat here is that a lack of relationship between the later position of these epithelia and teeth does not mean that these tissues do not influence tooth position. It may well be that, at some early critical moment, an ecto–endo–dermal boundary provides positional information or orientation for some teeth. Many early signalling centres in the body (such as the apical ectodermal ridge for limb development) are known to disassemble once they have done their job⁸.

Nonetheless, Soukup and colleagues' study removes the basis for theories depending on 'co-option' processes that would require migration of epithelial cells, and redirects future research. We need to study the molecular co-option of tooth or denticle genetic programs, a process that might have occurred several times independently in the history of jawed vertebrates. Which gene-regulatory regions are involved in switching on key regulators of tooth or denticle initiation in both epithelial and mesenchymal tissue? How, where and when did these genomic regions evolve? Are the same regions driving expression in ectoderm and endoderm? Are the regions involved in patterning denticle fields also used for organizing feathers and hair? And where are the 'atoms of information' that initiate, position and shape a tooth or denticle, and make its internal structure different from that of a dermal bone?

One can expect that there are combinations of transcription-factor proteins, bound

to unique tooth or denticle genetic regulatory elements, that drive the earliest molecular inducers of teeth or denticles. Such combinations, and their phylogenetic distribution and history, will be the ultimate arbiters of palaeontological arguments over dental and denticle homology⁹.

Discovery of the underlying tooth- or denticle-forming molecular programs will require transgenic analyses in paddlefish, catfish and sharks. Such analyses are not yet possible; nor are in-depth reconstructions of gene-regulatory sequences or bound transcription factors, necessitating the development of new experimental and bioinformatics approaches. Cracking such hard technical nuts will require strong intellectual teeth as well as robust

body armour, given the vigour of opinion on this subject.

Georgy Koentges is at the Systems Biology Centre, University of Warwick, Coventry CV4 7AL, UK.

e-mail: g.koentges@warwick.ac.uk

1. Soukup, V., Epperlein, H.-H., Horáček, I. & Cerný, R. *Nature* **455**, 795–798 (2008).
2. Sellman, S. *Odontologist Tidskrift* **54**, 1–28 (1946).
3. Thesleff, I. *Am. J. Med. Genet.* **140A**, 2530–2535 (2006).
4. Janvier, P. *Early Vertebrates* (Oxford Sci. Publ., 1996).
5. Johanson, Z. & Smith, M. M. *Biol. Rev.* **80**, 303–345 (2005).
6. Young, G. C. J. *Vert. Paleontol.* **23**, 987–990 (2003).
7. Sobkow, L., Epperlein, H.-H., Herklotz, S., Straube, W. L. & Tanaka, E. M. *Dev. Biol.* **290**, 386–397 (2006).
8. Guo, Q., Loomis, C. & Joyner, A. L. *Dev. Biol.* **264**, 166–178 (2003).
9. Koentges, G. *Nature* **451**, 658–663 (2008).

GEOMORPHOLOGY

How Tibet might keep its edge

Lewis A. Owen

The stability of the margins of the Himalayan–Tibetan mountain belt constitutes a puzzle. Repeated damming of major Tibetan rivers by glaciers, so controlling river erosion, is a possible explanation.

The collision of the Indian and Asian continental plates is the most dramatic tectonic event that Earth has experienced in the past 50 million years. It resulted in the formation of the Himalayan–Tibetan mountain belt, the growth of which initiated the south Asian monsoon, created some of the world's greatest rivers and gorges, and established the most highly glaciated realm outside polar regions. The combi-

nation of high topography, monsoon climate and great rivers and glaciers has produced intense erosion at the margins of these mountains, bringing deeply buried rocks quickly to the surface, most rapidly at the western and southeastern edges¹. Defining landscape development in the Himalaya and Tibet, and the factors involved, is among the greatest challenges facing geoscientists, and it is one tackled



Figure 1 | River deep, mountain high. At this site, which is just above where the Yarlung Tsangpo River (just off the photograph to the right) enters its gorge and slices its way through the Himalaya, impressive moraines extend from the flanks of the Namche Barwa massif and mark the limit of glaciation. The location corresponds to the place just upstream of the knick point shown in Figure 2.

B. HALLET

would have to occur to explain the presence of denticles on the outer covering of sharks and other more basal vertebrates^{5,6}.

Both theories hinge on the idea that there is an inherent difference in the inductive power of ectoderm and endoderm, and that migration of one or the other is the crucial factor in tooth formation. Implicit in this is the notion that tooth and denticle anatomy reflects embryonic origins — that is, that actual tooth or denticle histology can reveal which embryonic tissue was the key source.

Soukup *et al.*¹ now provide experimental grounds to debunk such ideas by testing the spatial distribution of ectoderm and endoderm in relation to erupting teeth. They took advantage of a line of transgenic axolotls⁷ — a group of amphibians that have ample teeth in their mouth — in which cell lineages can be fluorescently labelled. By grafting ectoderm or endoderm labelled with a marker known as green fluorescent protein from transgenic into normal axolotls, and vice versa, the authors show that there is no relationship between ectodermal and endodermal origin and the shape or nature of the resulting teeth — at least at the point when such teeth become visible. The enamel of teeth can be of ectodermal, endodermal or mixed origin. This is a dramatic finding. It means that one cannot infer relative distributions of ectoderm and endoderm from tooth or denticle anatomy even in a living species, let alone in a fossil.

The caveat here is that a lack of relationship between the later position of these epithelia and teeth does not mean that these tissues do not influence tooth position. It may well be that, at some early critical moment, an ecto–endo–dermal boundary provides positional information or orientation for some teeth. Many early signalling centres in the body (such as the apical ectodermal ridge for limb development) are known to disassemble once they have done their job⁸.

Nonetheless, Soukup and colleagues' study removes the basis for theories depending on 'co-option' processes that would require migration of epithelial cells, and redirects future research. We need to study the molecular co-option of tooth or denticle genetic programs, a process that might have occurred several times independently in the history of jawed vertebrates. Which gene-regulatory regions are involved in switching on key regulators of tooth or denticle initiation in both epithelial and mesenchymal tissue? How, where and when did these genomic regions evolve? Are the same regions driving expression in ectoderm and endoderm? Are the regions involved in patterning denticle fields also used for organizing feathers and hair? And where are the 'atoms of information' that initiate, position and shape a tooth or denticle, and make its internal structure different from that of a dermal bone?

One can expect that there are combinations of transcription-factor proteins, bound

to unique tooth or denticle genetic regulatory elements, that drive the earliest molecular inducers of teeth or denticles. Such combinations, and their phylogenetic distribution and history, will be the ultimate arbiters of palaeontological arguments over dental and denticle homology⁹.

Discovery of the underlying tooth- or denticle-forming molecular programs will require transgenic analyses in paddlefish, catfish and sharks. Such analyses are not yet possible; nor are in-depth reconstructions of gene-regulatory sequences or bound transcription factors, necessitating the development of new experimental and bioinformatics approaches. Cracking such hard technical nuts will require strong intellectual teeth as well as robust

body armour, given the vigour of opinion on this subject.

Georgy Koentges is at the Systems Biology Centre, University of Warwick, Coventry CV4 7AL, UK.

e-mail: g.koentges@warwick.ac.uk

1. Soukup, V., Epperlein, H.-H., Horáček, I. & Cerný, R. *Nature* **455**, 795–798 (2008).
2. Sellman, S. *Odontologist Tidskrift* **54**, 1–28 (1946).
3. Thesleff, I. *Am. J. Med. Genet.* **140A**, 2530–2535 (2006).
4. Janvier, P. *Early Vertebrates* (Oxford Sci. Publ., 1996).
5. Johanson, Z. & Smith, M. M. *Biol. Rev.* **80**, 303–345 (2005).
6. Young, G. C. J. *Vert. Paleontol.* **23**, 987–990 (2003).
7. Sobkow, L., Epperlein, H.-H., Herklotz, S., Straube, W. L. & Tanaka, E. M. *Dev. Biol.* **290**, 386–397 (2006).
8. Guo, Q., Loomis, C. & Joyner, A. L. *Dev. Biol.* **264**, 166–178 (2003).
9. Koentges, G. *Nature* **451**, 658–663 (2008).

GEOMORPHOLOGY

How Tibet might keep its edge

Lewis A. Owen

The stability of the margins of the Himalayan–Tibetan mountain belt constitutes a puzzle. Repeated damming of major Tibetan rivers by glaciers, so controlling river erosion, is a possible explanation.

The collision of the Indian and Asian continental plates is the most dramatic tectonic event that Earth has experienced in the past 50 million years. It resulted in the formation of the Himalayan–Tibetan mountain belt, the growth of which initiated the south Asian monsoon, created some of the world's greatest rivers and gorges, and established the most highly glaciated realm outside polar regions. The combi-

nation of high topography, monsoon climate and great rivers and glaciers has produced intense erosion at the margins of these mountains, bringing deeply buried rocks quickly to the surface, most rapidly at the western and southeastern edges¹. Defining landscape development in the Himalaya and Tibet, and the factors involved, is among the greatest challenges facing geoscientists, and it is one tackled



Figure 1 | River deep, mountain high. At this site, which is just above where the Yarlung Tsangpo River (just off the photograph to the right) enters its gorge and slices its way through the Himalaya, impressive moraines extend from the flanks of the Namche Barwa massif and mark the limit of glaciation. The location corresponds to the place just upstream of the knick point shown in Figure 2.

B. HALLET

by Korup and Montgomery on page 786 of this issue². From their studies, centred on southeast Tibet, they propose that there are previously unrecognized links between glaciation, erosion and bedrock uplift.

There are various possible explanations of the interactions between climate, erosion and uplift in the Himalaya and Tibet, specifically for the western (Nanga Parbat) and southeastern (Namche Barwa) massifs. These include the 'tectonic aneurysm' model^{1,3}, according to which localized feedback between focused river erosion and rock uplift leads to dynamic interactions between topographic stresses, thermal weakening of the crust and uppermost mantle, and deformation. In another model, the 'glacial buzz-saw' model⁴, glacial erosion controls the elevation of mountains regardless of uplift rates. These ideas have stimulated much research and debate, particularly over the relative importance of rivers and glaciers in driving landscape development.

Korup and Montgomery² add a new aspect to these debates. They argue that glaciation profoundly influences river erosion and sedimentation by damming water courses, which in turn controls the spatial distribution of erosion and the rate at which bedrock is brought to the Earth's surface. These interactions effectively help preserve the margin of Tibet, which would otherwise have retreated.

The authors' study focuses on the Yarlung Tsangpo River and its tributaries that drain Namche Barwa and form one of the deepest gorges on Earth (Fig. 1; see also the map on page 786 of this issue). These rivers, and others in similar high-mountain plateaux, should progressively erode towards their heads, resulting in steady degradation of a plateau's margins⁵. This is manifested in the headward migration of a river's 'knick point', an abrupt change in gradient in the longitudinal profile of a river. Such migration lowers the upstream elevation of the channel and basin; consequently, knick points should decline over time, and become less apparent as erosion progresses.

However, the margin of Tibet at Namche Barwa has apparently been stable over at least the past million years⁶, essentially preserving the knick points along the Yarlung Tsangpo and its tributaries. As Korup and Montgomery point out, the stability of the entire Himalaya–Tibetan margin poses a conundrum. Previous explanations have invoked the existence of a dynamic equilibrium between tectonic uplift and erosion⁷, or the role of landslides in protecting valley floors and in impeding channel incision⁸.

Korup and Montgomery² used a combination of remote sensing and field investigation. They specifically looked for moraines — huge deposits of rocks and debris that pile up in front and at the sides of glaciers, and remain when a glacier retreats. They thereby identified 260 locations where glaciers have dammed the Yarlung Tsangpo and its tributaries above the heads of knick points. They also observed that

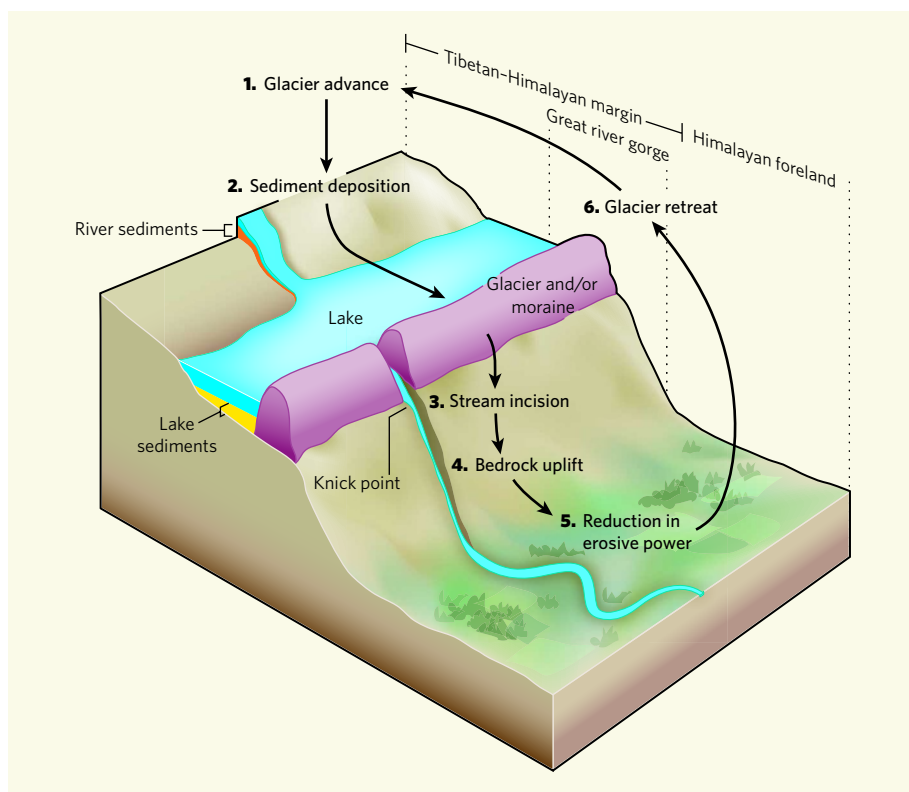


Figure 2 | Possible mechanism for the preservation of the edge of southeast Tibet. The following is the sequence of events proposed by Korup and Montgomery² (diagram not to scale). **1**, Glacier advance dams a river. **2**, The deposition of sediments in the resulting upstream lake and stream both protects the valley floor from erosion and reduces the stream's erosive power. **3**, Streams draining from the glacier and/or moraine are confined, so stream flow is strong and downstream erosion increases. **4**, Stream incision into the bedrock essentially weakens the crust and enhances bedrock uplift to the surface. **5**, As streams travel farther from the knick point, smaller gradients reduce their erosive power. **6**, The glacier retreats, but the events are repeated numerous times as glaciers advance and retreat in response to climate change. The repetition of this mechanism maintains equilibrium between erosion and bedrock uplift, essentially preserving the topography of the margin of southeast Tibet.

broad river valleys with substantial terraces and thick sediments are present upstream of the moraines. In contrast, downstream of the moraines, stream channels are more confined and are usually entrenched in bedrock and sediments that were washed out from landforms produced by the glaciers. The overall result — protection of landforms from erosion upstream of a knick point, and increased erosion but uplift of bedrock downstream — is outlined in Figure 2.

The geological evidence presented by Korup and Montgomery² shows that glaciers in the Namche Barwa region have advanced and retreated numerous times during the past 20,000 years or so, probably oscillating in a similar fashion since at least the beginning of the Quaternary ice age (some 2.5 million years ago). The authors quantified the degree of glaciation in the region by reconstructing the altitudes of former equilibrium lines for glaciers, a proxy for the extent of glaciation and hence climate. In this way also, they show that damming of the Yarlung Tsangpo and its tributaries has been a recurrent process, and has probably exercised strong control over the development of the landscape on million-year timescales. This repeated damming probably

maintained rapid exhumation of bedrock at knick points, like that proposed in the tectonic-aneurysm model.

Korup and Montgomery's explanation of the interactions between glaciation, erosion and uplift in Namche Barwa is persuasive and will prompt fresh thinking among geoscientists. But how widely applicable might the explanation be? To answer this question, the generality and full significance of Korup and Montgomery's proposal will need to be assessed throughout the Himalaya and Tibet, and in other mountain belts.

Lewis A. Owen is in the Department of Geology, University of Cincinnati, Cincinnati, Ohio 45221, USA.
e-mail: lewis.owen@uc.edu

1. Zeitler, P. K. *et al.* *GSA Today* **11** (1), 4–9 (2001).
2. Korup, O. & Montgomery, D. R. *Nature* **455**, 786–789 (2008).
3. Koons, P. O. *et al.* *Am. J. Sci.* **302**, 749–773 (2002).
4. Brozović, N., Burbank, D. W. & Meigs, A. J. *Science* **276**, 571–574 (1997).
5. Finlayson, D. P., Montgomery, D. R. & Hallet, B. *Geology* **30**, 219–222 (2002).
6. Finnegan, N. J. *et al.* *Geol. Soc. Am. Bull.* **120**, 142–155 (2008).
7. Lavé, J. & Avouac, J. P. *J. Geophys. Res.* **106**, 26561–26591 (2001).
8. Ouimet, W. B. *et al.* *Geol. Soc. Am. Bull.* **119**, 1462–1476 (2007).

Malaria research in the post-genomic era

Elizabeth Ann Winzeler¹

For many pathogens the availability of genome sequence, permitting genome-dependent methods of research, can partially substitute for powerful forward genetic methods (genome-independent) that have advanced model organism research for decades. In 2002 the genome sequence of *Plasmodium falciparum*, the parasite causing the most severe type of human malaria, was completed, eliminating many of the barriers to performing state-of-the-art molecular biological research on malaria parasites. Although new, licensed therapies may not yet have resulted from genome-dependent experiments, they have produced a wealth of new observations about the basic biology of malaria parasites, and it is likely that these will eventually lead to new therapeutic approaches. This review will focus on the basic research discoveries that have depended, in part, on the availability of the *Plasmodium* genome sequences.

Over the millennia malaria has shaped the human genome, driving an evolutionary response whereby otherwise disadvantageous alleles have persisted in the 'at-risk' population as a means of providing protection from this deadly disease. In addition, it has negatively affected human society by decreasing productivity and economic growth, and increasing poverty¹. Although eradicated in some parts of the world, at present 40% of the world's population lives in malarial regions². On the other hand, renewed interest in malaria prevention, the use of insecticide-impregnated bed nets and more efficacious drugs may have reduced the number of cases in the last several years, and there have been sporadic reports that some areas are seeing fewer malaria cases³.

The estimated 515 million cases of human malaria² each year are generally caused by four species, including *Plasmodium falciparum*, *P. vivax*, *P. ovale* and *P. malariae*, and are transmitted by the bites of female anopheline mosquitoes (Fig. 1). *P. falciparum* malaria has the greatest toll on human health, primarily in children under five. In addition to the characteristic fever and anaemia, neurological involvement may lead to an unarousable coma and in some cases death. If a child survives severe malaria his or her ability to learn may be impaired⁴, reducing his or her lifelong potential. *P. falciparum* is found at high levels in Africa, whereas in Asia and the Americas it is more common to find *P. vivax* malaria, which produces fewer fatalities but which nevertheless can be severe⁵. Whereas *P. falciparum* malaria is associated with a recent exposure to an infected mosquito, *P. vivax* and *P. ovale* parasites enter human hosts via the bite of an infected mosquito and then may remain dormant in the liver as the hypnozoite form for months or years, producing no outward manifestations of disease but resulting in relapses months or years after an individual has left a malaria endemic region. This is probably an adaptation to a temperate or subtropical climate where mosquitoes may not be present throughout the year.

Malaria control

Most malaria blood infections can be effectively treated with existing drugs, although resistance is a serious problem (reviewed in ref. 6). Chloroquine and antifolate drugs such as the sulphadoxine-pyrimethamine combination are safe and are still used in some regions because of their low cost. However, their continued use poses considerable risk due to widespread parasite resistance to these drugs as well as to others. At present the World Health Organization is recommending

the use of drugs that contain artemisinins—antiprotozoal endoperoxides from sweet wormwood (*Artemisia annua*)—in areas such as sub-Saharan Africa where resistance is common. Although artemisinin-based drugs have many attractive features including potency and rapid action, they are rapidly eliminated and have a complex chemical structure, which has thwarted researchers' attempts to synthesize them

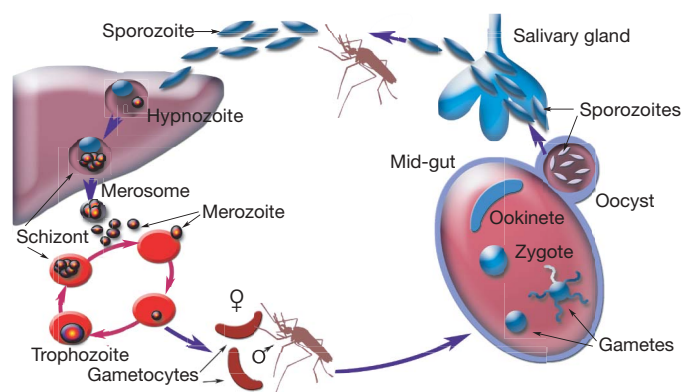


Figure 1 | Diagram of the malaria parasite's life cycle. Malaria is transmitted by the bite of a mosquito in which hundreds of sporozoites are released into the vertebrate host's bloodstream. The parasites eventually migrate to the liver—passing through some cell types such as Kupfer cells—and form parasitophorous vacuoles in hepatocytes. At this stage they can either remain dormant as a hypnozoite form (*P. vivax* or *P. ovale*), or initiate development that results in the production of thousands of merozoites. The parasites then induce detachment of the infected hepatocyte, allowing it to migrate to the liver sinusoid where budding of parasite-filled vesicles called merozoites⁷⁶ occurs. The new merozoites quickly invade erythrocytes where they replicate, sometimes synchronously, in a cycle that may correspond to the cycle of fever and chills in malaria. In response to a cue that is not well understood, some parasites differentiate into male and female gametocytes, which are the forms taken up by the mosquito and which can live quiescently in the bloodstream for weeks. Once they enter the mosquito via a blood meal they rapidly undergo transition into activated male and female gametes. The motile and short-lived diploid parasite form, the ookinete, migrates out of the blood meal, across the peritrophic matrix to the mid-gut wall where an oocyst is formed. After a meiotic reduction in chromosome number sporozoites are formed within the oocyst. Eventually the oocyst ruptures and the sporozoites migrate to the salivary gland where they await transfer to the vertebrate host.

¹Genomics Institute of the Novartis Research Foundation, San Diego, California and The Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

inexpensively in the laboratory. On the other hand, compounds with similar pharmacophores have been synthesized⁷ and it may even be possible to engineer yeast to produce such compounds⁸. Currently the drug of last resort, parasite resistance to artemisinin-based drugs has been observed in the rodent models for malaria⁹, and there are a few potential cases of treatment failure in patients¹⁰. Although these reports have not yet been confirmed in cultured parasites, it seems likely that full-blown resistance to artemisinins will eventually be observed, emphasizing the need for continued drug discovery research. Primaquine remains the only licensed drug that can provide a radical cure of *P. vivax* but this drug, in addition to its other liabilities, produces haemolytic anaemia in individuals with glucose-6-phosphate deficiency.

Although drugs currently work well against malaria, a vaccine that targets *P. vivax* or *P. falciparum* would probably reduce much of the poverty associated with malaria and greatly assist in complete malaria eradication. Recently, encouraging results have been obtained for a vaccine, named 'RTS,S', from GlaxoSmithKline Biologicals that is targeted for licensing in 2011. The vaccine contains a recombinant malaria protein fused to the surface of a hepatitis B protein and is based on the *P. falciparum* circumsporozoite protein (CSP), an abundant surface protein associated with the pre-erythrocytic phases of parasite development. In limited clinical trials the vaccine has been shown to reduce the number of severe cases of malaria¹¹ and delay the time to the first clinical episode. Unfortunately, the same studies also showed that children still developed malaria, raising questions about efficacy and drawing comparisons to recent unsuccessful and costly attempts to develop an HIV vaccine. Despite this, the data show that RTS,S is likely to decrease malaria severity and morbidity—were RTS,S to fail, there are still more than 40 subunit vaccines in development and 16 in clinical trials (<http://www.malariavaccine.org> or reviewed in ref. 12).

It has been known for a number of years that if sporozoites—the parasite form found in mosquito salivary glands—are attenuated and given as a vaccine, protection superior to that observed with a natural infection is acquired. This has led to the development of a whole-organism vaccine¹³. Although this vaccine model may be less practical than single-subunit vaccines owing to the cost, distribution challenges and problems with achieving consistently good manufacturing protocols, it still remains the most effective method for inducing protective immunity and efforts are being made to overcome these obstacles. Although the fact that we still don't have a fully protective and licensed malaria vaccine despite decades of effort may be disheartening, one must keep in mind that the development of drugs and vaccines, which are just now entering large-scale clinical trials, was initiated decades ago in a very different era for malaria research. In this period researchers were limited by their ability to work with malaria parasites in the laboratory, and thus vaccine research efforts were mostly focused on a relatively small number of very abundant proteins that could be easily studied.

Malaria genomes

In 2002 the complete genome sequence of *P. falciparum* was published¹⁴, along with a partial sequence of *P. yoelii yoelii*, a rodent parasite often used in vaccine research¹⁵. In 2005 partial sequences of several other rodent parasites were published¹⁶, and in this issue the genome sequences of *P. vivax* (a human malaria parasite) and *P. knowlesi* (primarily a monkey parasite) are described^{17,18}. In addition, the human genome sequence as well as several races (PEST, M and S) of *Anopheles gambiae*, the species of mosquito that transmits *P. falciparum* malaria, have been determined (<http://genome.wustl.edu/> and <http://www.tigr.org/>) and in some cases published¹⁹. Genome sequences promised a new paradigm in which drug targets might be plucked from the genome²⁰, causing an acceleration of preclinical candidates into the drug and vaccine pipelines. Although it may be too early to determine whether public health will be improved, knowledge of the genome has markedly enhanced the pace of basic research on parasite biology, which may eventually lead

to the production of new drugs and vaccines. Here we review how genome sequencing has benefited malaria research and opened new avenues of research that may eventually lead to new cures.

The *Plasmodium* genomes^{14–18} are estimated to contain 23–27 million bases, 14 chromosomes and ~5,500 genes, including many members of multigene families likely to be associated with immune evasion and antigenic variation. They are also rich in low-complexity regions and most have a high A+T content. However, the genome of *P. falciparum* is exceptionally (A+T)-rich (79.6%) throughout the genome, whereas only portions of the genome of *P. vivax* show similar levels of richness (overall A+T content of 67.7%), indicating that the extreme A+T content in itself has probably not too much to do with the disease. Nevertheless, it is possible that genes that are exceptionally (A+T)-rich in both species may be more recombinogenic and more likely to be involved in immune evasion. Indeed, associations between (A+T)-richness and recombination frequency have been observed in many species. There are differences between the species. For example, in *P. falciparum* many of the multigene families involved in immune evasion are, for the most part, located near the ends of chromosomes and are often transcriptionally silent, suggesting possible mechanisms by which genetic diversity could be generated by non-homologous recombination across chromosomes and epigenetic regulation could be established. In contrast, members of multigene families in *P. knowlesi* are scattered across the chromosomes and are not strictly subtelomeric¹⁸. Despite the differences, 77% of the proteins are conserved across the different species.

Although parasite genome sequences are available, many of the genes identified do not have homologues in commonly studied model organisms and often lack a clear cellular function¹⁴. Even with the recently completed *P. vivax* study almost one-half of the predicted genes still lack characterized orthologues in other systems. Considering the impact that malaria parasites have had on human health it seems remarkable that we know as little as we do. For example, we do not know the basis for sex determination in malaria parasites or exactly how parasites become committed to sexual development. Little is known about the liver stages and how the parasites home to the liver but then pass through transverse some cells while forming parasitophorous vacuoles in others. The situation is particularly acute in *P. vivax*, which cannot be cultured. In fact there is debate as to whether the hypnozoite, the dormant parasite liver form involved in relapse in *P. vivax* malaria, is a distinct stage or merely an early exoerythrocytic form whose development has been arrested. Recently it was shown that parasite metabolism inside a human may differ from parasite metabolism in laboratory culture, again highlighting how little is known about the role of the parasite genes during active patient infections²¹. Although research into such questions may be regarded as peripheral to the central objective of finding better drugs and vaccines, further investigation into basic biology will probably assist translational research in serendipitous ways.

Functional studies of malaria genomes

The goal of functional genomics is to determine what the different genes comprising the genome are likely to do for the organism (in this case the *Plasmodium* species), often using high-throughput tools such as microarrays or mass spectrometers. Transcriptional and proteomic analyses of the complement of genes within the genome at different life cycle stages can serve as a form of 'virtual genetics' in which regulons of co-transcribed genes can be defined by virtue of their expression pattern across many different conditions. Thus, if a gene shows a large induction during early liver stage development, there is a good chance that this is the time when its protein product will be required^{22,23}. These data can then be used to predict possible functions for parasite genes and to overcome the lack of powerful forward genetic methods that have advanced knowledge in so many other organisms. Comprehensive expression analyses of different parasite life cycle stages have been performed for a number of malaria parasite species^{16,21,23–25}. Because these data are only predictions, other

high-throughput methods for collecting genome-wide data can assist in determining which of the genes revealed in a gene expression experiment are likely to be involved in a process of interest. Two-hybrid mapping²⁶ and proteomics^{16,27,28} produce complementary data. Comprehensive analysis of the parasite proteome from male and female gametocytes has revealed genes contributing to the phenotypic differences between parasites of different sexes²⁹. Genome database websites such as PlasmoDB (<http://www.plasmodb.org>) have served to disseminate the data from such studies and have assisted the research activities of everyone interested in malaria. New genetic tools^{30,31} have been introduced that assist those interested in testing these predictions and may eventually be used in large-scale knockout or mutagenesis projects. Genome-scale protein crystallization projects are also underway, which should also help the pace of drug development^{32,33}.

One piece of information that has come from comprehensive proteomic and transcriptional profiling is that translational control probably has an important role in gene regulation during development. Although experiments have generally shown that there is a good correlation between transcript and protein abundance, there are notable discrepancies. In a number of cases genes appear to be transcribed but then not translated until the organism has made the rapid transition between warm- and cold-blooded hosts. Most notably, transcripts needed for gamete formation in the mosquito are produced in gametocytes in the mammalian host, but only translated after the transition—initiated by an increase in xanthurenic acid and a calcium-dependent signalling cascade³⁴—has been made^{16,35}. Specific genes involved in translational repression have been characterized³⁶. Although translational silencing clearly has a role in gene regulation when the parasite moves from human to mosquito, it seems likely that this is also occurring in the sporozoites, which inhabit the salivary glands of mosquitoes. The sporozoite form must also be prepared for a rapid and unexpected transition to a warm-blooded organism. Indeed, the liver-stage-associated protein 2 is not detected until 2 days after the parasite invades the liver³⁷, even though its cognate transcript is one of the most abundant in the sporozoite salivary gland transcriptome²⁴. Accordingly, it was noted that the malaria parasite genome seems to encode many genes involved in translational regulation, such as helicases and RNA binding proteins, but relatively few transcription factors¹⁴, and many of these are transcribed in gametocytes and sporozoites. Many non-coding RNAs are transcribed by the parasite genome and some of these may have interesting roles in controlling developmental processes³⁸.

Expression profiling has also revealed groups of genes with a probable involvement in the parasite's interaction with the mosquito²³ and which could be candidates for transmission-blocking vaccines. Transmission-blocking vaccines (reviewed in ref. 39), which would be directed towards proteins associated with the sexual phases of development (gametocytes or gametes), would not protect from the clinical manifestations of malaria but would altruistically prevent an infected individual from passing the disease on to the next person, potentially assisting in eradication. Functional genomic methods and mosquito genome sequence have been used to discover genes involved in immunity⁴⁰ and in host–parasite interactions. Work has shown that some of these genes are critical for parasite survival or are associated with the mosquito's ability to clear parasites⁴¹, suggesting new vector-based control strategies⁴². Although there are of course logistical, environmental and ethical issues associated with releasing recombinant mosquitoes that are unable to transmit malaria, the idea remains appealing.

Genome-wide analysis of antigenic variation

Creating catalogues of genes potentially involved in parasite processes or understanding parasite development may seem arcane and unrelated to human health; however, genome-wide transcription studies can shed light on how parasites evade the host immune system. It is generally accepted that transcription of genes known to be involved in antigenic variation can switch among different members

of a gene family and that this may contribute to immune evasion and pathogenicity⁴³. Expression analysis and genome sequence has permitted the transcription of each of the 60 *var* genes, which encode versions of the variant surface antigen, PfEMP1, to be monitored, ultimately providing clues as to how antigenic variation in parasites may be regulated^{44–46}. Genome-wide transcriptional and other studies have indicated that transcriptional switching may not be confined to *var* genes, encoding PfEMP1 proteins, and involve a large number of other different gene families such as those involved in parasite invasion^{47,48}.

Notably, these expression data have also helped in the development of innovative vaccination strategies. For example, analysis of high-throughput *P. falciparum* proteomic data²⁷ revealed one exceptionally abundant sporozoite protein⁴⁹, subsequently named CelTos in *P. berghei*⁵⁰, which was more immunogenic than some of the historical antigens such as CSP. Antigens derived from this protein are found in an experimental vaccine that is entering clinical trials⁵¹. It has also been known that vaccination with radiation-attenuated parasites can lead to better protective immunity than can be achieved with a natural infection. Expression analysis has led to the identification of parasite genes that are specifically transcribed while the parasite resides in the mosquito salivary gland⁵². Disruption of these genes^{53,54}, or ones with similar patterns of transcription⁵⁵, has led to genetically attenuated parasites which are unable to colonize successfully the vertebrate host but which nevertheless induce a protective immune response.

Genetic diversity in malaria parasites

The identity and geographic distribution of alleles that contribute to drug resistance forms the foundation for public policy on antimalarial drugs. Genome-dependent methods have the potential to change how genes involved in drug resistance are discovered. In the past genes involved in drug resistance were identified through mapping studies⁵⁶, or through candidate gene approaches where alleles predicted to have a role in drug resistance in a different organism were correlated with resistance. Because microsatellite typing has previously revealed that haplotypes that surround the alleles involved in chloroquine resistance⁵⁷ and antifolate resistance⁵⁸ are statistically over-represented in parasite populations, there has been interest in collecting single-nucleotide polymorphism data in order to create a 'haplotype' map that can be used to identify new regions of the genome in disequilibrium in different drug-selected parasite populations^{59–61}. Genome-dependent methods have also revealed new candidate genes that may be involved in drug resistance. Comparative genome hybridization with high-density oligonucleotide arrays show that different strains of *P. falciparum* contain numerous copy number variations (CNVs) that may be associated with drug resistance. For example, most strains carry diverse CNVs centred around the multidrug resistance transporter as well as around the first gene in the folate pathway, GTP cyclohydrolase⁶². This latter gene amplification event may be a signature of widespread antifolate drug use.

Studies of genetic variation^{59–63} have provided clues as to why a universally effective single-subunit malaria vaccine may be difficult to develop. Clearly there are vastly different rates of variability in different parasite gene classes, with many genes involved in 'house-keeping' functions, such as ribosomal proteins, DNA replication enzymes, or components of the cytoskeleton, exhibiting very low levels of variability as one compares the genomes of different *P. falciparum* isolates from different continents, suggesting a recent population bottleneck or a selective sweep⁶⁴. In contrast, many genes that are members of the multigene families, or which are membrane proteins, show very high levels of variability⁶⁵, possibly indicating a basal evolution rate that is much higher than for housekeeping genes, occurring even during standard laboratory culture. These data also show high levels of allelic diversity in many of the genes that correspond to antigens used in vaccine trials, including CSP, the apical membrane antigen 1 and the merozoite surface protein 1. In some

cases the high level of variability in these proteins may explain allele-specific immunity that is sometimes observed⁶⁶. These diversity data suggest that it may be difficult to develop a single-subunit vaccine that can provide universal protection, but they also point to new vaccine candidates that show imprints of host immune selection.

From the genome to cell biology

One of the more exciting recent developments that has resulted from the genome sequencing efforts has been the observation that many of the proteins that eventually find their way out of the parasite and onto the surface of infected red blood cells carry a conserved protein export motif, called a VTS or PEXEL motif^{67,68}. After a cell is infected, malaria parasites form a parasitophorous vacuole in the host cell and soon afterwards a tubovesicular network to promote protein trafficking (Fig. 2). Trafficking is likely to go in both directions, with both import of nutrients and export of proteins involved in immune evasion occurring. Remarkably, the PEXEL/VTS motif is also found in exported proteins from the plant pathogen, *Phytophthora infestans*, the organism responsible for the blight that caused the Irish potato famine⁶⁹. Here, the motif is attached to small proteins introduced into the plant cytoplasm where they interfere with the plant defence systems by preventing protease activation and subsequent apoptosis^{70,71}.

Many of the 400 or so *P. falciparum* proteins that contain the PEXEL/VTS motif seem to be involved in creating 'knobby' structures on the surface of the infected red blood cell and are also associated with cytoadherence and antigenic variation; however, the motif is also found in proteins of parasites that do not cytoadhere, such as *P. vivax*⁷². In addition, whereas many are transcribed at the time that the tubovesicular network is established during development in the erythrocyte, a few seem to be transcribed in the pre-erythrocytic phase, at about the same time that proteins required for early liver stage development are transcribed⁷². One such protein is CSP, the target

of the RTS,S vaccine, which bears two PEXEL/VTS motifs⁷³. CSP has attracted attention over the years because it is probably the most abundant protein synthesized in the infectious sporozoite stage of the parasite's life cycle. The protein is immunogenic, functioning potentially as the immunodominant antigen in the irradiated sporozoite vaccine⁷⁴. Although it has an essential role in the mosquito phases⁷⁵ of parasite development, its exact cellular role has been a matter of debate. Recently the PEXEL/VTS motif from CSP was attached to reporter proteins, where it directed the reporter protein into the cytoplasm of cells⁷³, raising the possibility that CSP might have a role in host–pathogen interactions. Notably, transient overexpression of CSP in HeLa cells seems to have a substantial effect on host transcription, downregulating many genes involved in immune signalling and upregulating other genes involved in cell adhesion and possibly apoptosis⁷³. It has been noted that parasites can grow for many days in the liver without triggering apoptosis or an inflammatory response, and once they have matured in the hepatocytes and are ready to be released the infected hepatocytes 'round up' and lose their ability to adhere⁷⁶. Active immune modulation is known to occur in other parasitic infections, such as schistosomiasis, and is likely to be a signature of a successful pathogen. Indeed, a secreted protein kinase affecting gene expression has recently been identified as a major virulence factor in *Toxoplasma gondii*, a closely related apicomplexan parasite^{77,78}.

The genome and drug discovery

The availability of genome sequence has also coincided with a renaissance in malaria drug discovery, as malaria research has become more tractable and more funding has been directed towards drug development through the establishment of public–private partnerships. When the genome sequences first began to appear in the literature, researchers noted several pathways that appeared to be present in parasites but absent in higher eukaryotes²⁰. These included the type II fatty acid biosynthesis pathway that is associated with the apicoplast, a unique organelle found in apicomplexan parasites. Although new, parasite-specific drug targets are still of substantial interest, recent drug discovery campaigns may be shifting from the single-enzyme screening approaches to cell-based methods where one can test for inhibition of all essential proteins simultaneously^{79–81}. Cell-based methods have been less attractive in the past because of the difficulties associated with finding a compound's mechanism of action. The availability of genome-dependent methods (for example, tiling microarrays or deep full-genome sequencing that can pick up new mutations that have been introduced in the laboratory as a consequence of drug pressure), sensitive proteomic methods that can be used in affinity chromatography experiments, as well as new robots and high-content imaging systems may also make the daunting task of discovering what compounds are doing in the cell less challenging.

Despite the advances there is still much work ahead. For example, the RTS,S and irradiated sporozoite vaccines are both imperfect and many feel that other approaches should be examined in parallel, especially as RTS,S is only partially efficacious. There are also logistical hurdles associated with distributing either an attenuated or killed organism vaccine. Testing vaccines remains extremely expensive because there is no way to tell whether a vaccine is effective except through the use of human subjects. There are hurdles in the drug development front because it is likely that most of the compounds that have shown the ability to kill parasites in the laboratory setting will not be modifiable into something that could be feasibly and safely delivered to patients in Africa. In addition to problems with delivering drugs, educating local physicians and accessing at-risk populations, there are inadequate methods available for surveillance and to even determine whether the World Health Organization's roll back malaria campaigns (<http://www.rbm.who.int/>) are having an effect. Despite this, during this renaissance in malaria research optimism has reached a point where some are calling for new campaigns to

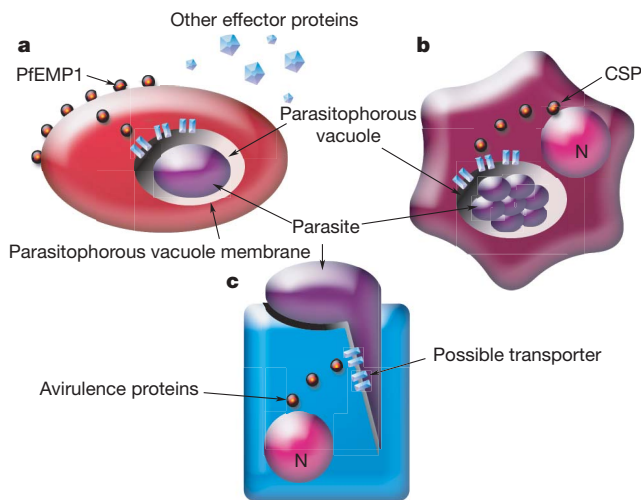


Figure 2 | Export pathways shared by eukaryotic pathogens. **a**, In erythrocytic stages of the life cycle the parasite establishes a tubovesicular network within the erythrocyte and then exports effector proteins (for example, PfEMP1s, Maurer's cleft proteins, FIKK kinases) containing a PEXEL/VTS motif across the parasitophorous vacuole (PV) into the cytoplasm of the infected erythrocyte^{67,68}. Some of these proteins, such as PfEMP1s, eventually reach the surface of the infected erythrocyte where they have a role in antigenic variation and immune evasion. Some may also be released in the bloodstream. **b**, In the liver stages of a malaria infection CSP and potentially other effector proteins may pass through the cytoplasm and eventually reach the nucleus (N). These proteins may interfere with the import of nuclear factor- κ B⁷³, a protein needed for activation of the human immune response, or they may have other roles controlling hepatocyte physiology. **c**, Small eukaryotic plant pathogens use a similar pathway with PEXEL/VTS-containing effector proteins released into the plant cytoplasm where they may interfere with plant innate immune defences⁸².

eliminate malaria completely. If basic research continues to be a priority and if support is sustained, new drugs and effective vaccines are likely to be developed, and this could make the goal of global malaria eradication achievable.

1. Sachs, J. & Malaney, P. The economic and social burden of malaria. *Nature* **415**, 680–685 (2002).
2. Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. & Hay, S. I. The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**, 214–217 (2005).
3. Okiro, E. A. *et al.* The decline in paediatric malaria admissions on the coast of Kenya. *Malar. J.* **6**, 151, doi:10.1186/1475-2875-6-151 (2007).
4. Boivin, M. J. *et al.* Cognitive impairment after cerebral malaria in children: a prospective study. *Pediatrics* **119**, e360–e366 (2007).
5. Tjitra, E. *et al.* Multidrug-resistant *Plasmodium vivax* associated with severe and fatal malaria: a prospective study in Papua, Indonesia. *PLoS Med.* **5**, e128 (2008).
6. Baird, J. K. Effectiveness of antimalarial drugs. *N. Engl. J. Med.* **352**, 1565–1577 (2005).
7. Vennerstrom, J. L. *et al.* Identification of an antimalarial synthetic trioxolane drug development candidate. *Nature* **430**, 900–904 (2004).
8. Ro, D. K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).
9. Afonso, A. *et al.* Malaria parasites can develop stable resistance to artemisinin but lack mutations in candidate genes *atp6* (encoding the sarcoplasmic and endoplasmic reticulum Ca^{2+} ATPase), *tctp*, *mdr1*, and *cg10*. *Antimicrob. Agents Chemother.* **50**, 480–489 (2006).
10. Wongsrichanalai, C. & Meshnick, S. R. Declining artesunate-mefloquine efficacy against falciparum malaria on the Cambodia-Thailand border. *Emerg. Infect. Dis.* **14**, 716–719 (2008).
11. Aponte, J. J. *et al.* Safety of the RTS,S/AS02D candidate malaria vaccine in infants living in a highly endemic area of Mozambique: a double blind randomised controlled phase I/IIb trial. *Lancet* **370**, 1543–1551 (2007).
12. Todryk, S. M. & Hill, A. V. S. Malaria vaccines: the stage we are at. *Nature Rev. Microbiol.* **5**, 487–489 (2007).
13. Epstein, J. E. *et al.* Safety and clinical outcome of experimental challenge of human volunteers with *Plasmodium falciparum*-infected mosquitoes: an update. *J. Infect. Dis.* **196**, 145–154 (2007).
14. Gardner, M. J. *et al.* Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
15. Carlton, J. M. *et al.* Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* **419**, 512–519 (2002).
16. Hall, N. *et al.* A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* **307**, 82–86 (2005).
17. Carlton, J. M. *et al.* Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* doi:10.1038/nature07327 (this issue).
18. Pain, A. *et al.* The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature* doi:10.1038/nature07306 (this issue).
19. Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**, 129–149 (2002).
20. Jomaa, H. *et al.* Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573–1576 (1999).
21. Daily, J. P. *et al.* Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature* **450**, 1091–1095 (2007).
22. Tarun, A. S. *et al.* A combined transcriptome and proteome survey of malaria parasite liver stages. *Proc. Natl Acad. Sci. USA* **105**, 305–310 (2008).
23. Zhou, Y. *et al.* Evidence-based annotation of the malaria parasite's genome using comparative expression profiling. *PLoS ONE* **3**, e1570 (2008).
24. Le Roch, K. G. *et al.* Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503–1508 (2003).
25. Bozdech, Z. *et al.* The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, E5 (2003).
26. LaCount, D. J. *et al.* A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature* **438**, 103–107 (2005).
27. Florens, L. *et al.* A proteomic view of the *Plasmodium falciparum* life cycle. *Nature* **419**, 520–526 (2002).
28. Lasender, E. *et al.* Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. *Nature* **419**, 537–542 (2002).
29. Khan, S. M. *et al.* Proteome analysis of separated male and female gametocytes reveals novel sex-specific *Plasmodium* biology. *Cell* **121**, 675–687 (2005).
30. Balu, B., Shoue, D. A., Fraser, M. J. & Adams, J. H. High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element piggyBac. *Proc. Natl Acad. Sci. USA* **102**, 16391–16396 (2005).
31. Nkrumah, L. J. *et al.* Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. *Nature Methods* **3**, 615–621 (2006).
32. Vedadi, M. *et al.* Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related apicomplexan organisms. *Mol. Biochem. Parasitol.* **151**, 100–110 (2007).
33. Mehlin, C. *et al.* Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. *Mol. Biochem. Parasitol.* **148**, 144–160 (2006).
34. Billker, O. *et al.* Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. *Cell* **117**, 503–514 (2004).
35. Le Roch, K. G. *et al.* Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res.* **14**, 2308–2318 (2004).
36. Mair, G. R. *et al.* Regulation of sexual development of *Plasmodium* by translational repression. *Science* **313**, 667–669 (2006).
37. Siau, A. *et al.* Temperature shift and host cell contact up-regulate sporozoite expression of *Plasmodium falciparum* genes involved in hepatocyte infection. *PLoS Pathog.* **4**, e1000121 (2008).
38. Mourier, T. *et al.* Genome-wide discovery and verification of novel structured RNAs in *Plasmodium falciparum*. *Genome Res.* **18**, 281–292 (2008).
39. Saul, A. Mosquito stage, transmission blocking vaccines for malaria. *Curr. Opin. Infect. Dis.* **20**, 476–481 (2007).
40. Waterhouse, R. M. *et al.* Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* **316**, 1738–1743 (2007).
41. Blandin, S. *et al.* Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* **116**, 661–670 (2004).
42. Alphey, L. *et al.* Malaria control with genetically manipulated insect vectors. *Science* **298**, 119–121 (2002).
43. Kyes, S. A., Kraemer, S. M. & Smith, J. D. Antigenic variation in *Plasmodium falciparum*: gene organization and regulation of the var multigene family. *Eukaryot. Cell* **6**, 1511–1520 (2007).
44. Duraisingh, M. T. *et al.* Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* **121**, 13–24 (2005).
45. Voss, T. S. *et al.* A var gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria. *Nature* **439**, 1004–1008 (2006).
46. Dzikowski, R., Frank, M. & Deitsch, K. Mutually exclusive expression of virulence genes by malaria parasites is regulated independently of antigen production. *PLoS Pathog.* **2**, e22 (2006).
47. Stubbs, J. *et al.* Molecular mechanism for switching of *P. falciparum* invasion pathways into human erythrocytes. *Science* **309**, 1384–1387 (2005).
- This paper demonstrates how malaria parasites can switch transcription between different members of multigene families potentially under selection pressure, contributing to their ability to evade host immune response.**
48. Cortes, A. *et al.* Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog.* **3**, e107 (2007).
49. Doolan, D. L. *et al.* Identification of *Plasmodium falciparum* antigens by antigenic analysis of genomic and proteomic data. *Proc. Natl Acad. Sci. USA* **100**, 9952–9957 (2003).
- This study shows that vaccine candidates potentially superior to historical vaccine candidates may be discovered by mining proteomic and transcriptional data.**
50. Kariu, T., Ishino, T., Yano, K., Chinzei, Y. & Yuda, M. CelTOS, a novel malarial protein that mediates transmission to mosquito and vertebrate hosts. *Mol. Microbiol.* **59**, 1369–1379 (2006).
51. Moran, M. *et al.* The Malaria Product Pipeline: Planning for the Future (The George Institute for International Health, 2007).
52. Matuschewski, K. *et al.* Infectivity-associated changes in the transcriptional repertoire of the malaria parasite sporozoite stage. *J. Biol. Chem.* **277**, 41948–41953 (2002).
53. Mueller, A. K. *et al.* *Plasmodium* liver stage developmental arrest by depletion of a protein at the parasite-host interface. *Proc. Natl Acad. Sci. USA* **102**, 3022–3027 (2005).
54. Mueller, A. K., Labaied, M., Kappe, S. H. & Matuschewski, K. Genetically modified *Plasmodium* parasites as a protective experimental malaria vaccine. *Nature* **433**, 164–167 (2005).
- This work demonstrates how genes identified by transcriptional profiling (using cDNA sequencing) can lead to innovative vaccination strategies for malaria.**
55. van Dijk, M. R. *et al.* Genetically attenuated, P36p-deficient malarial sporozoites induce protective immunity and apoptosis of infected liver cells. *Proc. Natl Acad. Sci. USA* **102**, 12194–12199 (2005).
56. Sidhu, A. B., Verdier-Pinard, D. & Fidock, D. A. Chloroquine resistance in *Plasmodium falciparum* malaria parasites conferred by pfcrt mutations. *Science* **298**, 210–213 (2002).
57. Wootton, J. C. *et al.* Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. *Nature* **418**, 320–323 (2002).
- This work demonstrates that loci involved in drug resistance can be mapped using parasite population studies, setting the stage for more extensive analyses of parasite phenotypes.**
58. Roper, C. *et al.* Intercontinental spread of pyrimethamine-resistant malaria. *Science* **305**, 1124 (2004).
59. Jeffares, D. C. *et al.* Genome variation and evolution of the malaria parasite *Plasmodium falciparum*. *Nature Genet.* **39**, 120–125 (2007).
60. Volkman, S. K. *et al.* A genome-wide map of diversity in *Plasmodium falciparum*. *Nature Genet.* **39**, 113–119 (2007).
61. Mu, J. *et al.* Genome-wide variation and identification of vaccine targets in the *Plasmodium falciparum* genome. *Nature Genet.* **39**, 126–130 (2007).
62. Kidgell, C. *et al.* A systematic map of genetic variation in *Plasmodium falciparum*. *PLoS Pathog.* **2**, e57 (2006).
- This is the first genome-wide study of diversity, which also reveals that copy number variations are widespread in malaria parasites.**

63. Carret, C. K. *et al.* Microarray-based comparative genomic analyses of the human malaria parasite *Plasmodium falciparum* using Affymetrix arrays. *Mol. Biochem. Parasitol.* **144**, 177–186 (2005).
64. Joy, D. A. *et al.* Early origin and recent expansion of *Plasmodium falciparum*. *Science* **300**, 318–321 (2003).
65. Volkman, S. K. *et al.* Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. *Science* **298**, 216–218 (2002).
66. Polley, S. D. *et al.* *Plasmodium falciparum* merozoite surface protein 3 is a target of allele-specific immunity and alleles are maintained by natural selection. *J. Infect. Dis.* **195**, 279–287 (2007).
67. Hiller, N. L. *et al.* A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**, 1934–1937 (2004).
68. Marti, M., Good, R. T., Rug, M., Knuepfer, E. & Cowman, A. F. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**, 1930–1933 (2004).
69. Bhattacharjee, S. *et al.* The malarial host-targeting signal is conserved in the Irish potato famine pathogen. *PLoS Pathog.* **2**, e50 (2006).
This is the first evidence that eukaryotic microbes share similar host trafficking signals.
70. Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T. & Kamoun, S. A. Kazal-like extracellular serine protease inhibitor from *Phytophthora infestans* targets the tomato pathogenesis-related protease P69B. *J. Biol. Chem.* **279**, 26370–26377 (2004).
71. Armstrong, M. R. *et al.* An ancestral oomycete locus contains late blight avirulence gene *Avr3a*, encoding a protein that is recognized in the host cytoplasm. *Proc. Natl Acad. Sci. USA* **102**, 7766–7771 (2005).
72. Sargeant, T. J. *et al.* Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol.* **7**, R12 (2006).
73. Singh, A. P. *et al.* *Plasmodium* circumsporozoite protein promotes the development of the liver stages of the parasite. *Cell* **131**, 492–504 (2007).
This paper shows that the abundant surface protein CSP is exported to the host cytoplasm where it may have an immunomodulatory role.
74. Kumar, K. A. *et al.* The circumsporozoite protein is an immunodominant protective antigen in irradiated sporozoites. *Nature* **444**, 937–940 (2006).
75. Menard, R. *et al.* Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes. *Nature* **385**, 336–340 (1997).
76. Sturm, A. *et al.* Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313**, 1287–1290 (2006).
This work describes how hepatocyte-derived merosomes signal to the host to both ensure migration of the parasite into the bloodstream and to protect themselves from host immunity.
77. Taylor, S. *et al.* A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii*. *Science* **314**, 1776–1780 (2006).
78. Saeij, J. P. *et al.* *Toxoplasma* co-opts host gene expression by injection of a polymorphic kinase homologue. *Nature* **445**, 324–327 (2007).
79. Baniecki, M. L., Wirth, D. F. & Clardy, J. High-throughput *Plasmodium falciparum* growth assay for malaria drug discovery. *Antimicrob. Agents Chemother.* **51**, 716–723 (2007).
80. Chong, C. R., Chen, X., Shi, L., Liu, J. O. & Sullivan, D. J. Jr. A clinical drug library screen identifies astemizole as an antimalarial agent. *Nature Chem. Biol.* **2**, 415–416 (2006).
81. Plouffe, D. *et al.* *In silico* activity profiling reveals the mechanism of action of antimalarials discovered in a high-throughput screen. *Proc. Natl Acad. Sci. USA* **105**, 9059–9064 (2008).
82. Whisson, S. C. *et al.* A translocation signal for delivery of oomycete effector proteins into host plant cells. *Nature* **450**, 115–118 (2007).

Acknowledgements E. A.W. is grateful for support from the Keck Foundation, the Wellcome Trust, Novartis and NIH RO1 AI059742.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Correspondence should be addressed to E.A.W. (winzeler@scripps.edu).

Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*

Jane M. Carlton^{1,2}, John H. Adams³, Joana C. Silva^{4,5}, Shelby L. Bidwell¹, Hernan Lorenzi¹, Elisabet Caler¹, Jonathan Crabtree^{1,5}, Samuel V. Angiuoli^{5,8}, Emilio F. Merino², Paolo Amedeo¹, Qin Cheng⁹, Richard M. R. Coulson¹⁰, Brendan S. Crabb^{11,12}, Hernando A. del Portillo^{13,14}, Kobby Essien^{15,16}, Tamara V. Feldblyum⁵, Carmen Fernandez-Becerra¹³, Paul R. Gilson¹¹, Amy H. Gueye¹⁷, Xiang Guo¹, Simon Kang'a², Taco W. A. Kooij¹⁸, Michael Korsinczyk^{9,19}, Esmeralda V.-S. Meyer²⁰, Vish Nene^{4,5}, Ian Paulsen^{1,21}, Owen White^{5,6}, Stuart A. Ralph²², Qinghu Ren¹, Tobias J. Sargeant^{11,23}, Steven L. Salzberg⁸, Christian J. Stoeckert¹⁵, Steven A. Sullivan², Marcio M. Yamamoto²⁴, Stephen L. Hoffman²⁵, Jennifer R. Wortman^{5,7}, Malcolm J. Gardner^{1†}, Mary R. Galinski¹⁹, John W. Barnwell²⁶ & Claire M. Fraser-Liggett^{5,7}

The human malaria parasite *Plasmodium vivax* is responsible for 25–40% of the ~515 million annual cases of malaria worldwide. Although seldom fatal, the parasite elicits severe and incapacitating clinical symptoms and often causes relapses months after a primary infection has cleared. Despite its importance as a major human pathogen, *P. vivax* is little studied because it cannot be propagated continuously in the laboratory except in non-human primates. We sequenced the genome of *P. vivax* to shed light on its distinctive biological features, and as a means to drive development of new drugs and vaccines. Here we describe the synteny and isochore structure of *P. vivax* chromosomes, and show that the parasite resembles other malaria parasites in gene content and metabolic potential, but possesses novel gene families and potential alternative invasion pathways not recognized previously. Completion of the *P. vivax* genome provides the scientific community with a valuable resource that can be used to advance investigation into this neglected species.

Plasmodium vivax is the major cause of malaria outside Africa, mainly afflicting Asia and the Americas¹. Principally a disease affecting poor people living on the margins of developing economies, *vivax* malaria traps many societies in a relentless cycle of poverty. Intermittent transmission makes protective immunity rare, and the disease strikes all ages. Repeated acute febrile episodes of debilitating intensity can occur for months. In children this can lead to life-long learning impairment, whereas incapacitation of adults has tremendous direct economic consequences through lost productivity and depletion of meagre financial reserves. Drug resistance in *P. vivax* is spreading, hindering management of clinical cases, and reports of severe pathology, including respiratory distress and coma, are challenging the description of *P. vivax* malaria as 'benign' (ref. 1).

Several biological characteristics underlie the distinct pathogenic and epidemiological nature of *vivax* malaria. In contrast to *P. falciparum*, *P. vivax* is only capable of infecting reticulocytes, causing

severe anaemia by dyserythropoiesis and destruction of infected and uninfected erythrocytes despite much lower parasitaemias. *P. vivax* cannot infect Duffy-blood-group-negative reticulocytes (a trait shared with the closely related monkey malaria parasite *P. knowlesi*), and is thus absent from West Africa where Duffy negativity predominates². Differences in *Anopheles* mosquito dynamics allow *P. vivax* transmission in temperate climates not tolerated by *P. falciparum*. In such regions *P. vivax* infects hepatocytes but may persist as dormant hypnozoites for months or years before initiating blood-stage infections (relapses) during another transmission season.

Because *P. vivax* kills infrequently and is not amenable to continuous *in vitro* culture, it has been relatively little studied in comparison to *P. falciparum*. The *P. vivax* genome sequence we report here, and comparative analyses with sequenced malaria parasites *P. falciparum*³, the rodent parasite *P. yoelii yoelii*^{4,5}, and the primate parasite *P. knowlesi*⁶ (an excellent model for *in vivo* studies of human

¹The Institute for Genomic Research/J. Craig Venter Institute, 9704 Medical Research Drive, Rockville, Maryland 20850, USA. ²Department of Medical Parasitology, New York University Langone Medical Center, 341 East 25th Street, New York, New York 10010, USA. ³Department of Global Health, College of Public Health, University of South Florida, 3720 Spectrum Boulevard, Suite 304, Tampa, Florida 33612, USA. ⁴Department of Microbiology and Immunology, ⁵Institute for Genome Sciences, ⁶Department of Epidemiology and Preventive Medicine, ⁷Department of Medicine, University of Maryland School of Medicine, 20 Penn Street, Baltimore, Maryland 21201, USA. ⁸Center for Bioinformatics and Computational Biology, University of Maryland, College Park, Maryland 20742, USA. ⁹Drug Resistance and Diagnostics, Australian Army Malaria Institute, Weary Dunlop Drive, Gallipoli Barracks, Enoggera, Queensland 4051, Australia. ¹⁰Microarray Group, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. ¹¹The Walter & Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3050, Australia. ¹²Burnet Institute, 85 Commercial Road, Melbourne, Victoria 3004, Australia. ¹³Barcelona Centre for International Health Research, Hospital Clinic/IDIBAPS, Universitat de Barcelona Roselló 132, 4a planta, 08036 Barcelona, Spain. ¹⁴Institució Catalana de Recerca i Estudis Avançats, Passeig Lluís Companys, 23 08010 Barcelona, Spain. ¹⁵Center for Bioinformatics and Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104, USA. ¹⁶Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA. ¹⁷Hood College, Frederick, Maryland 21701, USA. ¹⁸Department of Parasitology, Heidelberg University School of Medicine, Im Neuenheimer Feld 324, Heidelberg 69120, Germany. ¹⁹Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia. ²⁰Emory Vaccine Center, Yerkes National Primate Research Center and Department of Medicine, Division of Infectious Diseases, Emory University, Atlanta, Georgia 30329, USA. ²¹Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney, New South Wales 2109, Australia. ²²Department of Biochemistry & Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Victoria 3010, Australia. ²³Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia. ²⁴Departamento de Parasitologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, Av. Lineu Prestes 1374, São Paulo, São Paulo 05508-900, Brazil. ²⁵Sanaria Inc., 9800 Medical Center Drive, Rockville, Maryland 20850, USA. ²⁶Malaria Branch, Division of Parasitic Diseases, National Center for Zoonotic, Vector-borne and Enteric Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30341, USA. †Present address: Seattle Biomedical Research Center, 307 Westlake Avenue N., Suite 500, Seattle, Washington 98109-5219, USA.

malaria), provide important insights into the biology of this neglected parasite.

Genome sequencing and characteristics

The ~26.8-megabase (Mb) nuclear genome sequence of *P. vivax* (Salvador I) was sequenced by whole-genome shotgun methods to tenfold coverage, with targeted gap closure and finishing, and manual curation of automated annotation. Details of these and other methods are given in Supplementary Information. Large contigs totalling ~22.6 Mb were assigned to the 14 *P. vivax* chromosomes; ~4.3 Mb of small subtelomeric contigs remain unassigned due to their repetitive nature (Supplementary Table 1). *P. vivax* chromosomes are unique among human *Plasmodium* species in exhibiting a form of isochore structure⁷, with subtelomeric regions of low G+C content and chromosome internal regions of significantly higher G+C content. We finished the subtelomeric ends of several *P. vivax* chromosomes, allowing us to define their isochore boundaries (Fig. 1).

In many aspects, the genomes of mammalian *Plasmodium* species (*P. falciparum*, *P. knowlesi*, *P. vivax*, *P. y. yoelii*) are uniform, ranging from 23 to 27 Mb across 14 chromosomes, and comprising ~5,500 genes, most of which (~51%) contain at least one intron (Table 1). However, differences in nucleotide bias can be extreme (for example, *P. vivax* and *P. falciparum* average a G+C content of ~42.3% and ~19.4%, respectively), and a large gene family found in *P. y. yoelii* raised its gene count to ~5,880 (ref. 4). A remarkable 77% of genes are orthologous between the four species (Supplementary Fig. 1); almost one-half of these encode conserved hypothetical proteins of unknown function. In *P. falciparum*, the high incidence of tandem repeats and low complexity regions (LCRs) in proteins, especially antigens, has led researchers to propose that LCRs are involved in immune evasion mechanisms, such as antigen diversification⁸ and reducing the host's antibody response to critical epitopes by acting as a 'smokescreen'⁹. We found that LCRs tend to constitute a smaller proportion of *P. vivax* proteins on average (39%) than *P. falciparum* proteins (60%; Supplementary Fig. 2), and that LCR expansion partly accounts for the slightly larger size of *P. falciparum* proteins (Supplementary Table 2), but how this relates to differences in immune evasion mechanisms between *P. vivax* and *P. falciparum* is unclear.

Notwithstanding the recent functional characterization of the apicomplexan AP2 family of transcriptional regulators in *Plasmodium*¹⁰, the parasite seems to lack most of the standard eukaryotic transcriptional machinery, such as transcription-associated proteins (TAPs)¹¹, but is rich in regulatory sequences¹², fostering the idea that gene expression regulation in *Plasmodium* is complex and unusual. Our initial studies found no significant differences in the TAP repertoire between *P. falciparum*, *P. vivax* and *P. knowlesi*, indicating that transcriptional mechanisms are similar in all three species (Supplementary Table 3). Genes encoding messenger RNA stability proteins containing a CCCH-zinc finger were abundant in all three species, affirming the importance of post-transcriptional regulation in the control of gene expression across *Plasmodium*. A genome scan of *P. vivax* for known core promoter elements such as TATA and CAAT boxes identified some candidates, but many of them lacked positional specificity. Similarly, a search for novel promoter elements in regions upstream of ~1,800 mapped transcription start sites (5' UTRs), and for RNA binding elements in ~1,300 3' UTRs, also failed to produce convincing candidates (data not shown). To determine whether binding sites are conserved between *P. falciparum* and other primate *Plasmodium* species, we searched for over-represented nucleotide 'words' in regions upstream of clusters of potentially co-regulated genes conserved in *P. vivax*, *P. falciparum*, *P. knowlesi* and *P. y. yoelii* (Supplementary Information). Seven putative novel regulatory binding sites conserved across at least two species were identified (Supplementary Table 4). These binding sites were associated with core eukaryotic processes such as dephosphorylation and with parasite-specific functions such as cell

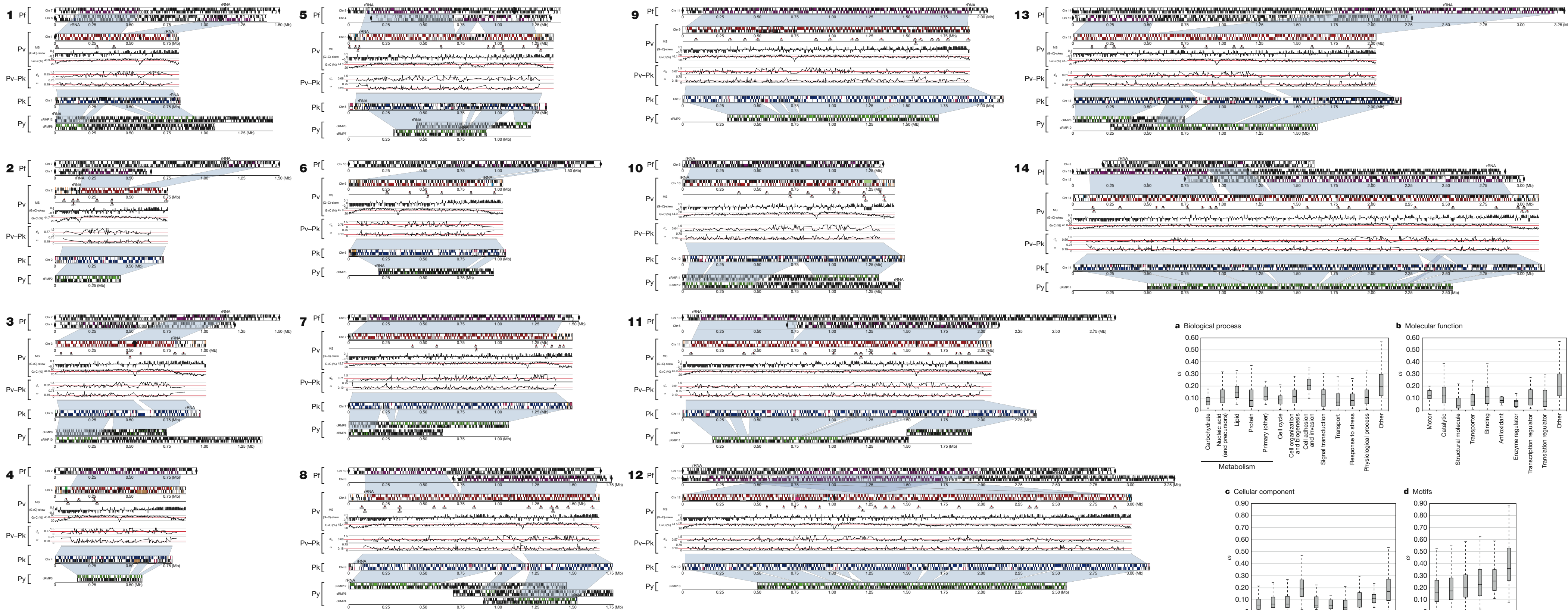
invasion. Independent support for two of our predicted sites comes from a recent report of the sporozoite-associated motif 5'-TGCAATGCA-3' and the merozoite invasion-related 5'-GTGTGCACAC-3' motif¹³. In our analysis these two sites, together with the dephosphorylation-associated motif 5'-GCACGCGTGC-3', were conserved across the four *Plasmodium* species.

Examination of natural parasite population structure is key to understanding transmission dynamics, the spread of drug resistance, and to design and test malaria control efforts. Many population studies have exploited the abundant polymorphic microsatellites in the *P. falciparum* genome, primarily simple sequence repeats such as [TA] dinucleotide and polyA/polyT¹⁴. We screened the *P. vivax* genome for microsatellites, identifying ~160 that are polymorphic between eight *P. vivax* laboratory lines (Fig. 1 and Supplementary Table 5). *P. vivax* microsatellites average 27.5% G+C, with an average repeat unit length of 3.1 nucleotides and an average copy number of 19.1. We found fewer microsatellites in *P. vivax* than in *P. falciparum* (as noted previously¹⁵), probably due to the more conventional nucleotide composition of the former. Even so, these genome-wide polymorphic markers are already facilitating studies of *P. vivax* population structure and genetic diversity^{16,17}.

Chromosome synteny and genome evolution

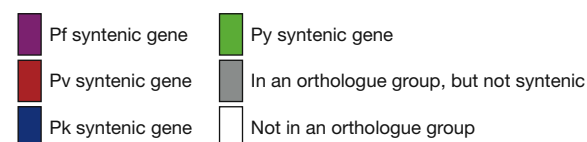
Previous studies have indicated significant conservation of gene synteny between *Plasmodium* parasites⁴ in direct proportion to their genetic distance. We generated a synteny map of *P. vivax*, *P. knowlesi*, *P. falciparum*, and the rodent malaria parasites *P. y. yoelii*, *P. berghei* and *P. chabaudi* (considered as a single lineage¹⁸ owing to their virtually complete synteny; Fig. 1). The *P. vivax* and *P. knowlesi* chromosomes are highly syntenic except for microsyntenic breaks at species-specific genes (in particular the *P. knowlesi* *kir* and *SICAvar* genes; see ref. 6); a previous study identified such breaks as foci for the evolution of host-parasite interaction genes¹⁸. The karyotypes of *P. vivax* and *P. knowlesi* correspond to the most parsimonious reconstruction of the ancestral form of the six species; the karyotypes of *P. falciparum* and the rodent malaria parasites can be reconstructed from this form through nine and six chromosomal rearrangements, respectively (Supplementary Fig. 3). No 'hotspots' of synteny breakage were identified, indicating that intersyntenic breakpoints were not 'reused' during the divergence of the species, and no obvious motifs except for (A+T)-rich regions and LCRs were identified in regions of the *P. vivax* genome predicted to have recombined to give

Figure 1 | Synteny maps showing the comparative organization of *Plasmodium* chromosomes. Putative orthologues were computed between *P. falciparum* (Pf), *P. vivax* (Pv), *P. knowlesi* (Pk) and *P. y. yoelii* (Py) proteomes and used to define blocks of synteny (shaded regions) between Py-Pk, Pv-Pk and Pv-Pf chromosomes. Genes on contigs that could not be assigned to chromosomes are not shown (see Supplementary Information). The composite rodent malaria parasite (cRMP) chromosomes generated in ref 18 are shown. Plots below the Pv chromosomes display the following: MS, the position of polymorphic microsatellites; (G+C)-skew, the base composition ((G-C)/(G+C)) within each strand in a 5-kb window; G+C, the percentage of bases in a 2-kb window that are G and C (%). Tracks within the Pv-Pk bracket show two evolutionary parameters: d_S (number of synonymous substitutions per synonymous site; top) and ω (ratio of d_N (non-synonymous substitutions per non-synonymous site) to d_S ; bottom). Horizontal red lines on the G+C (%), d_S and ω tracks indicate the average value for that chromosome. **a-d**, Distribution of selective constraints (ω) for the gene ontology classifications biological process (**a**), molecular function (**b**) and cellular component (**c**). Selective constraint is also shown for several motifs (**d**): proteins containing predicted transmembrane domains (TM) and/or signal peptides (SP); GPI-anchored proteins; and proteins predicted to be exported (exportome). Each grey box represents the interquartile range, which contains the sample's 25% to 75% range (quartiles Q1 to Q3, respectively), and the median is indicated (black horizontal bar within the interquartile range). Horizontal tick marks outside of the interquartile range show the range of all elements within $Q1 - 1.5 \times$ interquartile range and $Q3 + 1.5 \times$ interquartile range (~99.3% interval of a normal distribution).

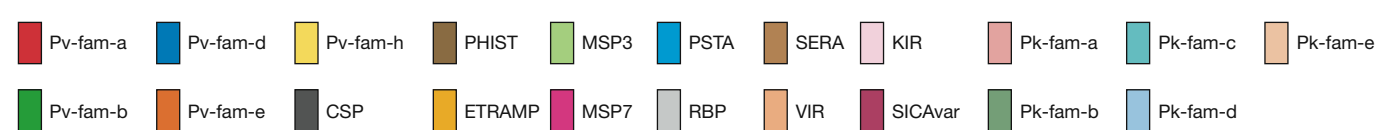


The Plasmodium genome

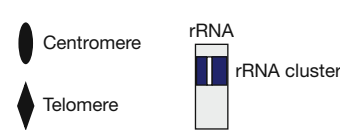
Conserved syntenic/orthologous groups



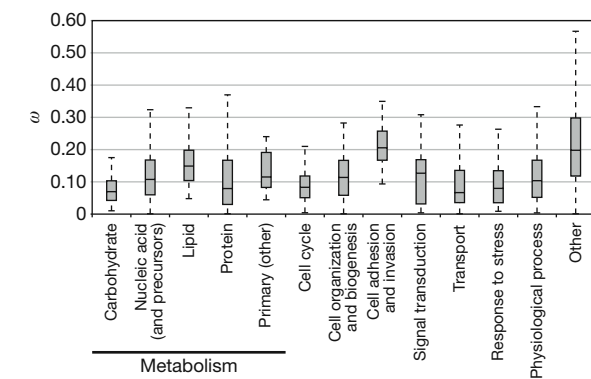
Gene families



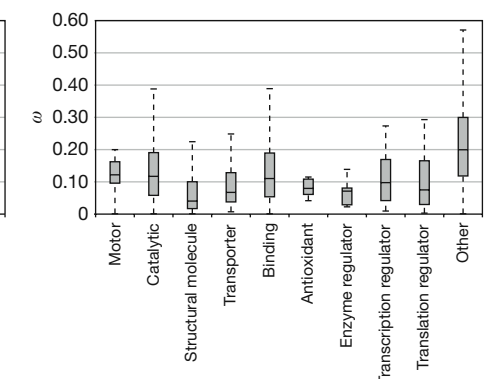
Other features



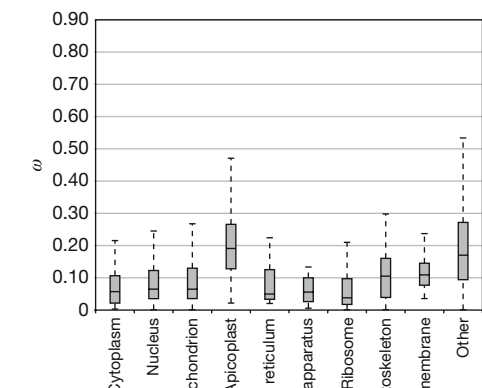
a Biological process



b Molecular function



c Cellular component



d Motifs

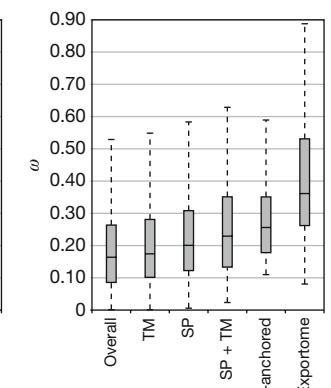


Table 1 | Comparison of nuclear genome features between four *Plasmodium* species

Feature	<i>P. vivax</i>	<i>P. knowlesi</i>	<i>P. falciparum</i>	<i>P. y. yoelii</i>
Genome				
Size (Mb)	26.8	23.5	23.3	23.1
Number of chromosomes	14	14	14	14
Coverage (fold)	10	8	14.5	5
G+C content (%)	42.3	37.5	19.4	22.6
Genes				
Number of genes	5,433*	5,188*	5,403*	5,878†
Mean gene length‡	2,164§	2,180	2,283	1,298†
Gene density (bp per gene)	4,462.9§	4,593	4,312	2,566†
Percentage coding‡	48.5§	47.4	52.6	50.6†
Genes with introns (%)	52.1§	51.6	53.9	54.2†
Exons				
Mean number per gene	2.5§	2.6	2.4	2.0†
G+C content (%)	46.5§	40.2	23.7	24.8†
Mean length (bp)	957§	836.8	935	641†
Introns				
G+C content (%)	49.8§	38.6	13.6	21.1†
Mean length (bp)	192§	224.4	179	209†
Intergenic regions				
G+C content (%)	42.5§	34.56	13.7	20.7†
Mean length (bp)	1,994§	2,049.4	1,745	859†
RNAs				
Number of tRNA genes	44	41	43	39
Number of 5S rRNA genes	3	0¶	3	3
Number of 5.8S/18S/28S rRNA units	7	5	7	4

* Including pseudogenes and partial genes, excluding non-coding RNA genes.

† Excluding partial genes.

‡ Excluding introns.

§ Excluding genes from 2,745 small (A+T)-rich contigs.

|| Excluding genes from 511 small contigs.

¶ Not present in *P. knowlesi* assembly version 4.0.

single *P. falciparum* chromosomes. Of the 3,336 orthologues between all six species, 3,305 (99%) were found to be positionally conserved (Supplementary Table 6).

We used 3,322 high-quality *P. vivax*/*P. knowlesi* orthologues to obtain maximum likelihood estimates of the rate of substitution at synonymous (d_s) and non-synonymous (d_N) sites, as well as ω (d_N/d_s ; Supplementary Table 7 and Fig. 1). *P. vivax* chromosomes differ significantly in their average values for both d_s and d_N , but the two variables are strongly correlated within and between chromosomes (Supplementary Fig. 4). The chromosomes also differ significantly in average %GC₄ (the G+C content in third codon positions of fourfold degenerate amino acids). This variable is positively correlated with average d_s and inversely correlated with chromosome length, such that synonymous sites in genes on the smallest chromosomes (~1 Mb) evolve ~1.5 times faster than genes on the two largest (~3 Mb) chromosomes (Supplementary Fig. 5). These observations strongly indicate the existence of heterogeneous mutation rates across the genome. It is unclear if this is due to cytosine-to-thymine deamination, which is more probable in (G+C)-rich regions, as it is not known whether DNA methylation occurs in *P. vivax*. The degree of selective constraint (ω) also varies across classes of genes. Genes encoding glycosylphosphatidylinositol (GPI)-anchored proteins, cell adhesion proteins, exportome proteins (proteins predicted to be exported) and proteins with transmembrane or signal peptide motifs, all of which are at least partly extracellular, were found to evolve significantly faster than genes involved in, for example, carbohydrate metabolism, enzyme regulation and cell structure (Fig. 1a–d and Supplementary Table 8). The host immune system, by targeting extracellular peptides, seems to have strongly influenced evolutionary rate variation between gene classes in *Plasmodium*.

A highly conserved *Plasmodium* metabolome

We found that key metabolic pathways, housekeeping functions and the repertoire of predicted membrane transporters are highly conserved between the *P. vivax* and *P. falciparum*³ proteomes

(Supplementary Table 9), suggesting that the two species have much the same metabolic potential. Conservation of metabolic processes also extends to the apicoplast, an apicomplexan plastid secondarily acquired from an ancient cyanobacterium. The apicoplast has lost photosynthetic function, but is essential to the parasite's metabolism, hosting nuclear-encoded proteins that are targeted to the apicoplast lumen by a conserved bipartite amino-terminal presequence. The complete genome sequence of *P. vivax* offers an opportunity to update and improve the apicoplast proteome that was predicted *in silico*³. Apicoplast-targeted proteins conserved in *P. vivax* participate in major metabolic processes previously recognized in *P. falciparum*¹⁹, such as complete type II fatty acid synthesis, isopentenyl diphosphate and iron sulphur cluster assembly pathways, and a fragmented haem synthesis pathway distributed between the apicoplast and mitochondria. Conservation of these pathways in *P. vivax* is important because synthetic pathways for type II fatty acid synthesis and isopentenyl diphosphates are targets for antimalarial chemotherapeutics²⁰. The revised *Plasmodium* apicoplast proteome (Supplementary Table 10) also clarifies the localization of two important processes. We show thiamine pyrophosphate biosynthesis, previously thought to take place in the apicoplast¹⁹, to be cytosolic. Conversely, we confirm a glyoxalase pathway in the apicoplast, with glyoxalase I and glyoxalase II enzymes being targeted there²¹; both enzymes are potential drug targets. Thus, comparison of overall apicoplast metabolic capabilities shows very few differences between *P. vivax* and *P. falciparum*.

P. vivax can form hypnozoites, a latent hepatic stage responsible for patent parasitaemia relapses months or even years after an initial mosquito-induced infection²². Hypnozoites survive most drugs that kill blood-stage parasites; complete elimination of *P. vivax* infections (radical cure) requires primaquine, the only licensed drug that can kill hypnozoite stages. However, resistance to the drug is spreading²³, and its use is contra-indicated in pregnant women or patients with glucose-6-phosphate dehydrogenase deficiency, which is common in malaria-endemic regions. After an initial examination of *P. vivax*-specific proteins failed to identify leads (Supplementary Table 11), we

some motifs at the amino terminus. Similarly to *P. falciparum* var genes, *in situ* hybridization analysis has shown that *P. vivax* chromosome ends localize to the nuclear periphery⁴⁰, where ectopic recombination favours the generation of variants and gene expansion. Although the repeat structure of *P. vivax* subtelomeric regions is not as extensive as that seen in *P. falciparum*⁶, *P. vivax* probably uses chromosomal exchange as a mechanism for generating antigenic diversity. VIR proteins represent an extremely diverse family, members of which currently appear more divergent than members of other partially characterized PIR families such as the *P. chabaudi* CIR (135 members) and the *P. berghei* BIR (245 members) families (Supplementary Fig. 9). Shared structural characteristics have been shown between VIR subfamily D proteins and the *P. falciparum* Pfmc-2tm family located at Maurer's clefts, and VIR subfamily A proteins and the *P. falciparum* SURFIN family found on the surface of infected erythrocytes⁴¹. We speculate that the extreme diversity and sub-structuring of VIR proteins indicate members' different subcellular localizations and functions, including immune evasion.

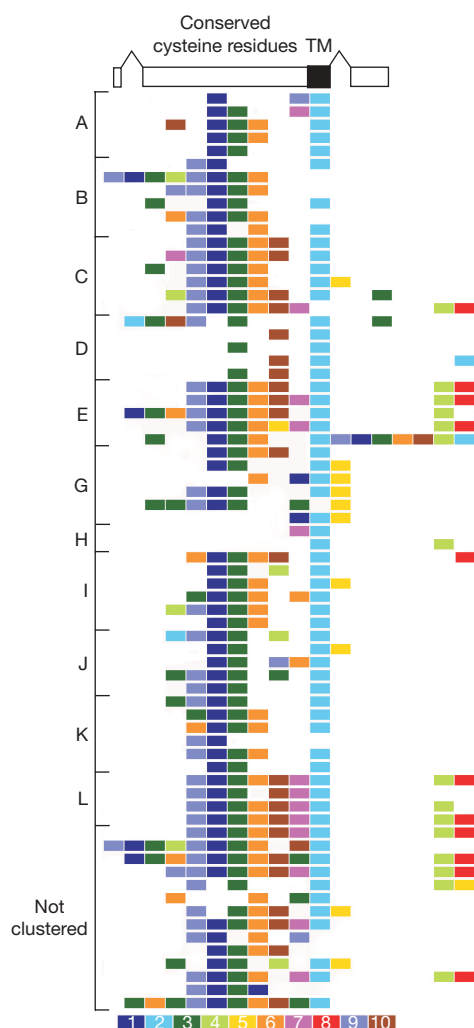


Figure 3 | VIR protein motifs and organization. The structure of an archetypal *vir* gene is shown at the top, followed by VIR motifs in each of the subfamilies A–E and G–L (subfamily F is not shown), arranged from the N terminus (left) to the C terminus (right). Consensus motif sequences numbered in decreasing order of statistical significance are shown colour coded below the figure. Motif 2, transmembrane (TM) domain; motif 3, PEXEL/VSP-like motif; all remaining motifs are predicted to be exposed globular domains. The overall organization and order of the motifs is maintained, with the central core motifs 9, 1, 3, 6 and 10 followed by C terminus motifs 7, 2, 4, 8 and 5 embedded in a variant-sized portion of the molecule. Motifs are listed in the Supplementary Information.

We identified eight novel gene families (Pv-fam-a to Pv-fam-e and Pv-fam-g to Pv-fam-i; Supplementary Table 16) in the *P. vivax* genome, most of which are located in subtelomeric regions (Fig. 1). Of particular interest are (1) the PvTRAG (Pv-fam-a) gene family (36 genes), one member of which was previously identified (it encodes a protein localized to the caveola-vesicle complex of infected erythrocytes, and has been shown to elicit a humoral immune response during the course of natural infections⁴²); and (2) the Pv-fam-e family (Supplementary Fig. 10), 36 copies of which are found in two loci on either side of the predicted centromere on chromosome 5, with one 10-gene locus present in a 47% G+C region, and a second 26-gene locus present in a 36% G+C region. Whereas *P. vivax* proteins have a fairly balanced codon composition, using all 61 sense codons almost equally (effective number of codons, $N_c = 54.2$), their orthologues in *P. falciparum* are more biased ($N_c = 37.5$), with G- and C-ending codons nearly absent from fourfold degenerate amino acids (Supplementary Table 17). However, *P. vivax* gene families, which are predominantly located in (A+T)-rich regions, have a codon composition of $N_c = 47$. This pattern suggests a strong influence of local mutation pattern on the nucleotide composition of genes and indicates a potential for differential gene expression.

Plasmodium drug interaction genes

The sexual stages of *P. vivax* are produced before the onset of clinical symptoms, permitting mosquito transmission early in an infection. Such early parasite transmission may delay development of resistance to many of the antimalarial drugs used to treat *vivax* malaria, despite the extensive long-term use of these drugs in regions endemic for both *P. vivax* and *P. falciparum*⁴³. Nevertheless, *P. vivax* can develop resistance to most of the current antimalarial drugs. To understand the interactions between antimalarial drugs and the parasite proteins implicated in drug binding and resistance, we examined crystal structures and developed homology models for several *P. vivax* proteins in the predicted proteome, and compared the predicted binding sites and reported mutations with those of their *P. falciparum* orthologues (Table 2).

Currently, the most efficacious novel antimalarial drugs are derivatives of artemisinin (qinghaosu) and atovaquone, used predominantly in combination therapies. Artemisinin derivatives, the most potent drugs recommended for treatment, may target a sarcoplasmic/endoplasmic reticulum Ca^{2+} ATPase (SERCA)-type protein, ATPase6 (ref. 44). We constructed homology models of *P. vivax* and *P. falciparum* ATPase6 and identified two residues in the putative active sites for artemisinin that differ between the two species (*P. vivax* A263 and S1008, equivalent to L263 and N1039 in *P. falciparum*). A change in residue 263 from leucine to alanine results in a threefold increase in susceptibility to artemisinin by *Xenopus laevis* expressing PfATPase6 (ref. 44), although the IC_{50} (concentration required to inhibit parasite growth by 50%) for some *P. vivax* field isolates appears higher than the IC_{50} for *P. falciparum*⁴⁵. However, it should be noted that clinical resistance of any human *Plasmodium* species to artemisinin derivatives has yet to be documented. Atovaquone, used in combination with the antifolate proguanil, selectively inhibits mitochondrial electron transport at the cytochrome *bc*(1) complex; mutations in the cytochrome *b* (*cytb*) gene can interfere with this inhibition, causing resistance. We constructed a homology model of *P. vivax* CYTB and compared it to the *P. falciparum* CYTB homology model⁴⁶, revealing almost identical structures, including the predicted atovaquone active sites. Although there are no reports of atovaquone treatment failures for *P. vivax* infections, our studies indicate that should resistance arise, the same sites in *P. vivax* CYTB may be implicated.

Towards a policy shift for *P. vivax* malaria

Despite the insights into parasite biology provided by the *P. vivax* genome, many important questions remain that can only be addressed by functional studies. For example, we were unable to find

Table 2 | Putative binding sites of *P. vivax* and *P. falciparum* orthologues potentially involved in artemisinin and atovaquone interactions.

Antimalarial drug	Gene	Gene PID*	Protein PID*	Homology (H) model	Drug binding site residues†	Polymorphisms associated or suspected with resistance‡	Conclusions
Artemisinin derivatives	<i>P. vivax</i> <i>atpase6</i>	60.7	70.0	H*	(I261, A263, F264, Q267, L268, I271, I275, A313, I942, I946, N949, I950, V953, A954, F957, S1008, L1009, I1010, L1015, Y1018, I1019)*	None reported	<i>P. vivax</i> predicted to be more susceptible to artemisinin than <i>P. falciparum</i>
	<i>P. falciparum</i> <i>atpase6</i>			H*	(I261, L263§, F264, Q267, L268, I271, I275, A313, I973, I977, N980, I981, V984, A985, F988, N1039, L1044, I1045, L1050, Y1053, I1054)*	E432K, A623E, S769N	
Atovaquone	<i>P. vivax</i> <i>cytb</i>	86.0	89.5	H*	(I119, F123, Y126, M133, V140, I141, L144, I258, P260, F264, F267, Y268, L271, V284, L285, L288)*	None reported	Resistance mutations in <i>P. falciparum</i> likely to affect the same residues in <i>P. vivax</i>
	<i>P. falciparum</i> <i>cytb</i>			H	I119, F123, Y126, M133, V140, I141, L144, I258, P260, F264, F267, Y268¶, L271, V284, L285, L288	M133I#, T142I#, L144F#, I258M#, F267I#, Y268S/N, L271V#, K272R#, P275T#, G280D#, V284K#	

The Supplementary Information contains an expanded table providing further details and references. *atpase6*, adenosine triphosphatase-6; *cytb*, cytochrome b; PID, percentage identity.

*Determined during this study.

†Underlined residues indicate polymorphism between *P. vivax* and *P. falciparum*.

‡Underlined polymorphisms indicate unique to *P. falciparum*.

§Predicted to be located in a drug active site and predicted to cause resistance.

|| Association with resistance not conclusive.

¶Predicted to be located in a drug active site and mutations known to cause resistance.

#Identified from *in vitro* drug selection studies.

differences in the predicted *P. vivax* proteome that might explain the rheological behaviour of *P. vivax*-infected erythrocytes, which remain flexible and can repeatedly pass through the spleen, unlike *P. falciparum*-infected reticulocyte, the rigidity of which facilitates cytoadherence and avoidance of splenic clearance⁴⁷. Studies of the hypnozoite transcriptome, although technically challenging, would radically increase our inadequate knowledge of the biology of this dormant form. Studies are currently underway to develop new *in vitro* culture systems⁴⁸, which could provide badly needed biological material for such functional studies.

The malaria research and control communities were challenged recently to establish once again the eradication of malaria as a policy goal⁴⁹. Given the significant contribution of *P. vivax* to the global malaria situation^{1,43}, it is imperative that these efforts include elimination of *P. vivax* as well as *P. falciparum*. Elimination of *P. vivax* presents special challenges, in particular the parasite's production of dormant hypnozoites that enables relapses long after the initial parasitaemia has cleared. Indeed, an important aspect of *P. vivax* eradication will be the development of new drugs to replace primaquine for radical cure. Although the development of new drugs targeting *P. vivax* liver stages is a formidable task, recent developments offer hope that this goal can be accomplished⁵⁰.

METHODS SUMMARY

Genome sequencing, assembly, mapping and annotation. *Saimiri boliviensis boliviensis* monkeys were infected with the Salvador I strain of *P. vivax* isolated from a patient from El Salvador. Extracted parasite DNA was used to make genomic DNA libraries for shotgun sequencing. Reads were assembled into scaffolds, inter-scaffold gaps closed, and scaffolds assigned to *P. vivax* chromosomes through hybridization of scaffold-specific probes to pulsed-field gel separated chromosomes. Gene prediction algorithms were used to predict gene models, and each model was manually checked for structural inconsistencies. Gene function was assigned using an automated annotation pipeline with subsequent manual curation.

Genome analysis. Methods for the *in silico* analysis of the genome sequence are described in the Supplementary Information.

Studies requiring laboratory experimentation. For polymorphic microsatellite identification, primers flanking 333 microsatellites identified from the genome sequence and designed for field studies where access to capillary electrophoresis equipment may not be possible were used to amplify the loci from eight worldwide *P. vivax* laboratory strains adapted to growth in monkeys (Brazil I, Miami II, Pakchong, Panama I, Nica, Thai II, Vietnam IV and Indonesia XIX). Amplicons were separated by electrophoresis on agarose gels and scored for size differences. For *vir* gene expression studies, cDNA was generated from total RNA

extracted from the Salvador I isolate and from three patient isolates from Brazil. Primers were designed to eight *vir* gene subfamilies and used to amplify the loci.

Received 18 January; accepted 8 August 2008.

- Price, R. N. et al. Vivax malaria: Neglected and not benign. *Am. J. Trop. Med. Hyg.* 77 (suppl. 6), 79–87 (2007).
- Miller, L. H., Mason, S. J., Clyde, D. F. & McGinniss, M. H. The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N. Engl. J. Med.* 295, 302–304 (1976).
- Gardner, M. J. et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511 (2002).
- Carlton, J. M. et al. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 419, 512–519 (2002).
- Hall, N. et al. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* 307, 82–86 (2005).
- Pain, A. et al. The genome of the simian and human malaria parasite *Plasmodium knowlesi*. *Nature* doi:10.1038/nature07306 (this issue).
- McCutchan, T. F., Dame, J. B., Miller, L. H. & Barnwell, J. Evolutionary relatedness of *Plasmodium* species as determined by the structure of DNA. *Science* 225, 808–811 (1984).
- Hughes, A. L. The evolution of amino acid repeat arrays in *Plasmodium* and other organisms. *J. Mol. Evol.* 59, 528–535 (2004).
- Anders, R. F. Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. *Parasite Immunol.* 8, 529–539 (1986).
- De Silva, E. K. et al. Specific DNA-binding by Apicomplexan AP2 transcription factors. *Proc. Natl Acad. Sci. USA* 105, 8393–8398 (2008).
- Coulson, R. M., Hall, N. & Ouzounis, C. A. Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Res.* 14, 1548–1554 (2004).
- van Noort, V. & Huynen, M. A. Combinatorial gene regulation in *Plasmodium falciparum*. *Trends Genet.* 22, 73–78 (2006).
- Young, J. A. et al. *In silico* discovery of transcription regulatory elements in *Plasmodium falciparum*. *BMC Genomics* 9, 70 (2008).
- Ferdig, M. T. & Su, X. Z. Microsatellite markers and genetic mapping in *Plasmodium falciparum*. *Parasitol. Today* 16, 307–312 (2000).
- Feng, X. et al. Single-nucleotide polymorphisms and genome diversity in *Plasmodium vivax*. *Proc. Natl Acad. Sci. USA* 100, 8502–8507 (2003).
- Imwong, M. et al. Relapses of *Plasmodium vivax* infection usually result from activation of heterologous hypnozoites. *J. Infect. Dis.* 195, 927–933 (2007).
- Joy, D. A. et al. Local adaptation and vector-mediated population structure in *Plasmodium vivax* malaria. *Mol. Biol. Evol.* 25, 1245–1252 (2008).
- Kooij, T. W. et al. A *Plasmodium* whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog.* 1, e44 (2005).
- Ralph, S. A. et al. Tropical infectious diseases: Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nature Rev. Microbiol.* 2, 203–216 (2004).
- Sato, S. & Wilson, R. J. The plastid of *Plasmodium* spp.: a target for inhibitors. *Curr. Top. Microbiol. Immunol.* 295, 251–273 (2005).
- Akoachere, M. et al. Characterization of the glyoxalases of the malarial parasite *Plasmodium falciparum* and comparison with their human counterparts. *Biol. Chem.* 386, 41–52 (2005).

22. Krotoski, W. A. The hypnozoite and malarial relapse. *Prog. Clin. Parasitol.* **1**, 1–19 (1989).
23. Baird, J. K. & Hoffman, S. L. Primaquine therapy for malaria. *Clin. Infect. Dis.* **39**, 1336–1345 (2004).
24. Sattabongkot, J. *et al.* Establishment of a human hepatocyte line that supports *in vitro* development of the exo-erythrocytic stages of the malaria parasites *Plasmodium falciparum* and *P. vivax*. *Am. J. Trop. Med. Hyg.* **74**, 708–715 (2006).
25. Fang, X. D., Kaslow, D. C., Adams, J. H. & Miller, L. H. Cloning of the *Plasmodium vivax* Duffy receptor. *Mol. Biochem. Parasitol.* **44**, 125–132 (1991).
26. Galinski, M. R., Medina, C. C., Ingravallo, P. & Barnwell, J. W. A. Reticulocyte-binding protein complex of *Plasmodium vivax* merozoites. *Cell* **69**, 1213–1226 (1992).
27. Cowman, A. F. & Crabb, B. S. Invasion of red blood cells by malaria parasites. *Cell* **124**, 755–766 (2006).
28. Gruner, A. C. *et al.* The Py235 proteins: glimpses into the versatility of a malaria multigene family. *Microbes Infect.* **6**, 864–873 (2004).
29. Duraisingh, M. T. *et al.* Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. *EMBO J.* **22**, 1047–1057 (2003).
30. Preiser, P. R., Jarra, W., Capod, T. & Snounou, G. A rhoptry-protein-associated mechanism of clonal phenotypic variation in rodent malaria. *Nature* **398**, 618–622 (1999).
31. Roussilhon, C. *et al.* Long-term clinical protection from falciparum malaria is strongly associated with IgG3 antibodies to merozoite surface protein 3. *PLoS Med.* **4**, e320 (2007).
32. Gilson, P. R. *et al.* Identification and stoichiometry of glycosylphosphatidylinositol-anchored membrane proteins of the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Proteomics* **5**, 1286–1299 (2006).
33. Sanders, P. R. *et al.* Distinct protein classes including novel merozoite surface antigens in Raft-like membranes of *Plasmodium falciparum*. *J. Biol. Chem.* **280**, 40169–40176 (2005).
34. Dzikowski, R., Templeton, T. J. & Deitsch, K. Variant antigen gene expression in malaria. *Cell. Microbiol.* **8**, 1371–1381 (2006).
35. del Portillo, H. A. *et al.* A superfamily of variant genes encoded in the subtelomeric region of *Plasmodium vivax*. *Nature* **410**, 839–842 (2001).
36. Fernandez-Becerra, C. *et al.* Variant proteins of *Plasmodium vivax* are not clonally expressed in natural infections. *Mol. Micro.* **58**, 648–658 (2005).
37. Marti, M., Good, R. T., Rug, M., Knuepfer, E. & Cowman, A. F. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**, 1930–1933 (2004).
38. Hiller, N. L. *et al.* A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**, 1934–1937 (2004).
39. Frank, M. *et al.* Strict pairing of *var* promoters and introns is required for *var* gene silencing in the malaria parasite *Plasmodium falciparum*. *J. Biol. Chem.* **281**, 9942–9952 (2006).
40. Scherf, A., Figueiredo, L. & Freitas-Junior, L. H. *Genomes and Molecular Cell Biology of Malaria Parasites* (Horizon Press, 2004).
41. Merino, E. F. *et al.* Multi-character population study of the *vir* subtelomeric multigene superfamily of *Plasmodium vivax*, a major human malaria parasite. *Mol. Biochem. Parasitol.* **149**, 10–16 (2006).
42. Jalah, R. *et al.* Identification, expression, localization and serological characterization of a tryptophan-rich antigen from the human malaria parasite *Plasmodium vivax*. *Mol. Biochem. Parasitol.* **142**, 158–169 (2005).
43. Mendis, K., Sina, B. J., Marchesini, P. & Carter, R. The neglected burden of *Plasmodium vivax* malaria. *Am. J. Trop. Med. Hyg.* **64**, 97–106 (2001).
44. Uhlemann, A. C. *et al.* A single amino acid residue can determine the sensitivity of SERCAs to artemisinins. *Nature Struct. Mol. Biol.* **12**, 628–629 (2005).
45. Russell, B. *et al.* Determinants of *in vitro* drug susceptibility testing of *Plasmodium vivax*. *Antimicrob. Agents Chemother.* **52**, 1040–1045 (2008).
46. Korsinczky, M. *et al.* Mutations in *Plasmodium falciparum* cytochrome *b* that are associated with atovaquone resistance are located at a putative drug-binding site. *Antimicrob. Agents Chemother.* **44**, 2100–2108 (2000).
47. Suwanarusk, R. *et al.* The deformability of red blood cells parasitized by *Plasmodium falciparum* and *P. vivax*. *J. Infect. Dis.* **189**, 190–194 (2004).
48. Udomsangpet, R., Kaneko, O., Chotivanich, K. & Sattabongkot, J. Cultivation of *Plasmodium vivax*. *Trends Parasitol.* **24**, 85–88 (2008).
49. Roberts, L. & Enserink, M. Malaria. Did they really say: eradication? *Science* **318**, 1544–1545 (2007).
50. Carraz, M. *et al.* A plant-derived morphinan as a novel lead compound active against malaria liver stages. *PLoS Med.* **3**, e513 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the *P. vivax* research community for their support, and in particular M. Gottlieb and V. McGovern for facilitating financial support. Funding came from the following sources: *P. vivax* sequencing, assembly and closure, US Department of Defense and National Institute of Allergy and Infectious Diseases; genome mapping, Burroughs Wellcome Fund; and selective constraint analysis, National Institute of General Medical Sciences. We wish to thank TIGR's SeqCore, Closure and IFX core facilities, E. Lee, J. Sundaram, J. Orvis, B. Haas and T. Creasy for engineering support, R. K. Smith Jr for annotation support, E. Lyons and H. Zhang for technical assistance, H. Potts for statistical analysis, T. McCutchan for rDNA sequence annotation, and S. Perkins for the *Plasmodium* phylogeny.

Author Information Sequence and annotation data for the genome are deposited in GenBank under the project accession number AAKM00000000 and are also available at the *Plasmodium* genome sequence database PlasmoDB (<http://plasmodb.org>). A minimal tiling path of clones covering each chromosome is available through the malaria repository MR4 (<http://www.mr4.org>), and a long-oligo array through the Pathogen Functional Genomics Resource Center (<http://pfgrc.jcvi.org>). Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.M.C. (jane.carlton@nyumc.org).

ARTICLES

SAP-controlled T–B cell interactions underlie germinal centre formation

Hai Qi^{1*}, Jennifer L. Cannons^{2*}, Frederick Klauschen¹, Pamela L. Schwartzberg² & Ronald N. Germain¹

Generation of long-term antibody-mediated immunity depends on the germinal centre reaction, which requires cooperation between antigen-specific T and B lymphocytes. In human X-linked lymphoproliferative disease and its gene-targeted mouse model, loss-of-function mutations in signalling lymphocyte activation molecule-associated protein (SAP, encoded by *SH2D1a*) cause a profound defect in germinal centre formation by an as yet unknown mechanism. Here, using two-photon intravital imaging, we show that SAP deficiency selectively impairs the ability of CD4⁺ T cells to stably interact with cognate B cells but not antigen-presenting dendritic cells. This selective defect results in a failure of antigen-specific B cells to receive adequate levels of contact-dependent T-cell help to expand normally, despite *Sap*^{−/−} T cells exhibiting the known characteristics of otherwise competent helper T cells. Furthermore, the lack of stable interactions with B cells renders *Sap*^{−/−} T cells unable to be efficiently recruited to and retained in a nascent germinal centre to sustain the germinal centre reaction. These results offer an explanation for the germinal centre defect due to SAP deficiency and provide new insights into the bi-directional communication between cognate T and B cells *in vivo*.

The germinal centre (GC) reaction, which supports antibody affinity maturation and the generation of B-cell memory^{1–3}, requires activation of CD4⁺ T cells by antigen-bearing dendritic cells (DCs), followed by differentiation of these T cells and their physical interactions with antigen-activated B cells⁴. The activated T cells promote the survival, proliferation and differentiation of the B cells by delivering contact-dependent helper signals in an antigen-specific fashion^{5–8}. Follicular helper T (T_{FH}) cells, a subset of activated CD4⁺ T cells that highly express the chemokine receptor CXCR5, the costimulatory molecules CD40L and ICOS, and SAP, are particularly important for a sustained GC response⁹.

In human X-linked lymphoproliferative disease and its gene-targeted mouse model, lack of functional SAP protein in T cells causes a profound defect in GC formation and humoral immunity^{10–17}, but how SAP regulates the GC-promoting T-helper cell response is unknown. An intracellular adaptor protein, SAP binds to signalling lymphocyte activation molecule (SLAM) and other transmembrane molecules of the SLAM family^{18,19}. It modulates T-cell antigen receptor signalling and promotes type 2 T helper (T_H2) cell differentiation^{13,14,20}. However, impaired T_H2 differentiation alone is not an adequate explanation for the profound GC defect seen in SAP-deficient animals^{17,21,22}. To better understand how SAP deficiency results in impaired GC formation, we took a combined approach of classical cellular immunology and intravital two-photon imaging, which allowed us to examine not only the activation phenotype of T cells but also the dynamics of their critical interactions with DCs and B cells *in vivo*²³.

Normal activation of *Sap*^{−/−} T cells by DCs

Because optimal T-cell activation and differentiation *in vivo* require long-lasting T–DC conjugation^{24–27}, we first examined the impact of SAP deficiency on T–DC interactions using OT-2 TCR transgenic T cells specific to I–A^b complexed with an ovalbumin-derived peptide (OVA₃₂₃). Behaviour of *Sap*^{−/−} and *Sap*^{+/+} T cells in the same draining lymph node were compared after co-transfer into mice previously

injected subcutaneously with OVA₃₂₃-pulsed DCs. *Sap*^{−/−} and *Sap*^{+/+} OT-2 T cells showed equivalent contact times with DCs presenting OVA₃₂₃ (Fig. 1a, b and Supplementary Movie 1). Consistent with this result, *Sap*^{−/−} and *Sap*^{+/+} OT-2 T cells proliferated and accumulated to a similar extent (Fig. 1c). Within the expanded pool of activated T cells of either genotype, similar frequencies of CXCR5⁺ICOS⁺ T_{FH} precursors were observed and both groups of T cells migrated into the B-cell follicular areas of the lymph node (Fig. 1d, e). These data indicate SAP deficiency does not grossly alter initial T-cell interaction with or activation by antigen-presenting DCs *in vivo*.

Selective defect in T–B cell interactions

During a T-dependent B-cell response, helper T cells activated by DCs must subsequently engage antigen-triggered B cells based on class II major histocompatibility complex (MHC class II)-restricted antigen recognition^{7,28}. This process is characterized by long-lasting, mobile conjugate pairs formed between cognate T cells and B cells²⁹. On the basis of these previous observations, we used intravital imaging to assess whether SAP deficiency affects antigen-dependent T–B cell interactions *in vivo*. *Sap*^{−/−} and *Sap*^{+/+} OT-2 T cells were co-transferred into mice together with wild-type B cells expressing the MD4 transgenic B-cell receptor, which recognizes hen egg lysozyme (HEL). As antigen, cross-linked conjugates of HEL and OVA (HEL–OVA) were used to allow activation of both T and B transgenic cells and to permit MHC-class-II-dependent antigen presentation to the OT-2 T cells by the MD4 B cells. After immunization, *Sap*^{+/+} OT-2 T cells engaged migrating MD4 B cells and formed long-lasting mobile conjugate pairs, whereas *Sap*^{−/−} OT-2 T cells showed predominantly brief contacts with the same cohort of MD4 B cells (Fig. 2a, b and Supplementary Movies 2 and 3; mean ± s.e.m. of the median T–B cell contact time by *Sap*^{+/+} versus *Sap*^{−/−} T cells: 19.2 ± 2.6 versus 8.2 ± 1.9 min, 5 experiments). Thus in notable contrast to the lack of significant effects on the duration of T–DC interactions (Fig. 1a, b),

¹Lymphocyte Biology Section, Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

²Genetic Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

*These authors contributed equally to this work.

the absence of SAP expression in T cells severely reduces the longevity of T–B cell interactions *in vivo*.

To examine more quantitatively the differential effect of SAP deficiency on T–antigen-presenting cell (APC) interactions, a flow-cytometry-based *in vitro* assay was used in which shear stress was imposed on unfixed conjugates between activated OT-2 T-cell blasts and OVA₃₂₃-pulsed DCs or activated B cells. Whereas a dose-dependent increase in T–B cell conjugate efficiency was observed with *Sap*^{+/+} OT-2 T cells, such conjugate formation by *Sap*^{-/-} T cells was low at all peptide concentrations tested (Fig. 2c, left panel). In contrast, within the same non-saturating peptide dose range, *Sap*^{-/-} T cells were as efficient as wild-type T cells in forming conjugates with DCs (Fig. 2c, right panel). To determine whether the defective adhesion to B cells by activated *Sap*^{-/-} T cells might have been programmed during their priming interactions with DCs and whether SAP expression is specifically required during T–B cell interactions, DNA constructs encoding either green fluorescent protein (GFP) or a GFP–SAP fusion protein were independently transfected into *Sap*^{-/-} T-cell blasts. Complementation with GFP–SAP but not GFP alone fully rescued T–B cell conjugate formation by transfected

Sap^{-/-} T-cell blasts (Fig. 2d; also see Supplementary Fig. 1). Thus, the requirement for SAP expression in T cells for optimal adhesion to B cells is contemporaneous with T–B cell interaction. SAP facilitates the recruitment of FYN kinase to cytoplasmic tails of SLAM-related proteins, which is important for SAP regulation of T_H2 cytokine production^{20,30} but less so for SAP-controlled GC development^{17,31}.

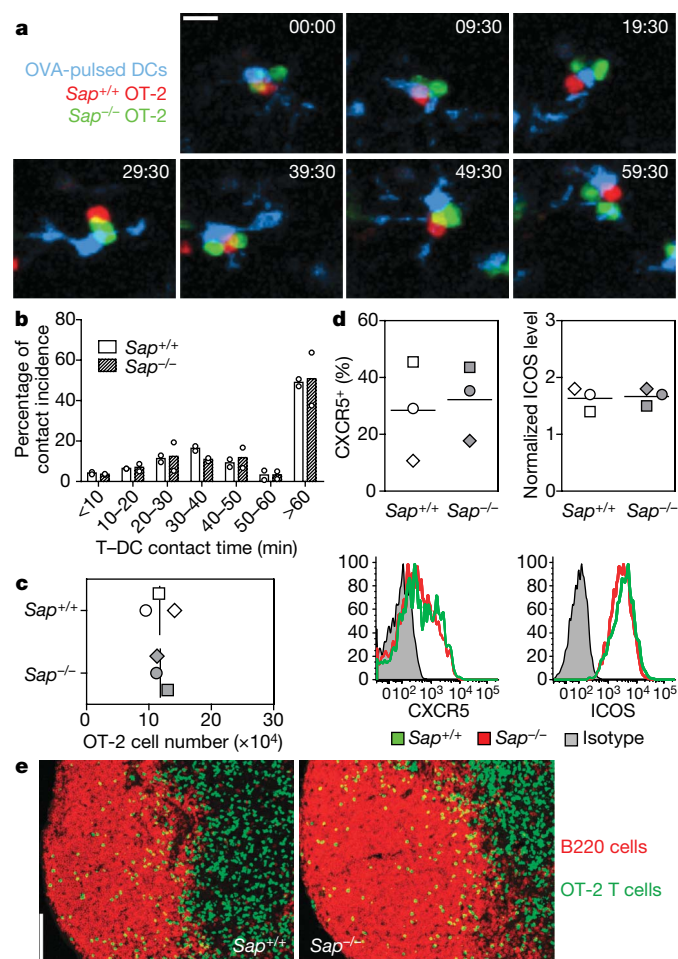


Figure 1 | *Sap*^{-/-} T cells normally interact with and are activated by DCs *in vivo*. **a**, Time-lapse images of *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells interacting with OVA₃₂₃-pulsed DCs *in vivo* (see also Supplementary Movie 1). Scale bar, 10 μm. Time is shown in min:s. **b**, Distribution of T–DC contact durations (circles, individual experiments; bars, means). A total of 232 and 224 contacts were scored for *Sap*^{+/+} and *Sap*^{-/-} T cells, respectively. **c–e**, T-cell activation phenotypes 96 h after transfer into recipients of OVA₃₂₃-pulsed DCs. Absolute numbers of OT-2 cells (**c**) and typical patterns of CXCR5 and ICOS expression (**d**). Individual symbols, mean of 3–4 mice per group per experiment; lines, mean of the 3 experiments. **e**, The distribution of activated T cells in B-cell follicles, representative of three experiments. Scale bar, 100 μm.

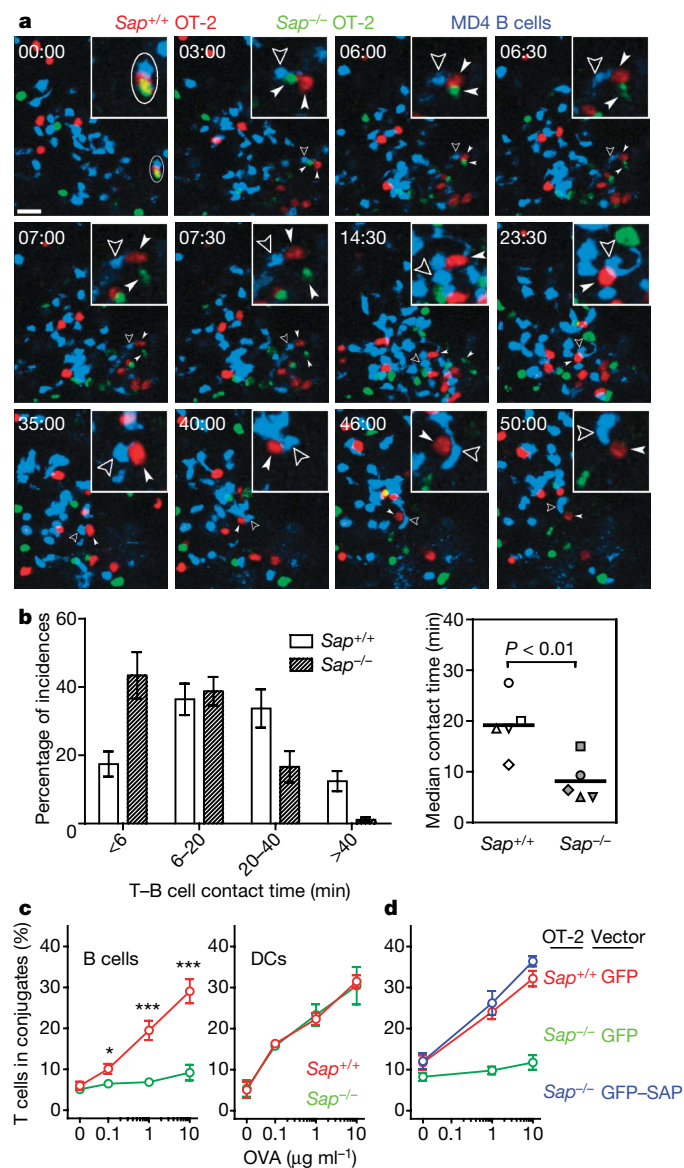


Figure 2 | *Sap*^{-/-} T cells are defective in adhesion to cognate B cells. **a**, Time-lapse images of *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells interacting with MD4 B cells 24–36 h after HEL–OVA immunization. Circle, a cell cluster composed of one *Sap*^{+/+} and one *Sap*^{-/-} T cell in contact with the same MD4 B cell at time zero. Solid arrowheads follow the two T cells, and open arrowheads highlight the B cell. Insets show the MD4 cell and its immediate surrounding (see also Supplementary Movie 2). Scale bar, 20 μm. **b**, Left, distribution of contact durations between T cells and MD4 B cells (mean ± s.e.m. of 5 experiments; 190 and 173 contacts scored for *Sap*^{+/+} and *Sap*^{-/-} T cells, respectively). Right, the median durations of T–B cell contacts measured in individual experiments. Line, mean of the 5 medians. **c**, Conjugation efficiency of *Sap*^{+/+} or *Sap*^{-/-} OT-2 T cells with OVA₃₂₃-pulsed B cells (left) or DCs (right), expressed as mean ± s.e.m. frequencies of CD4⁺CD19⁺ or CD4⁺CD11c⁺ conjugates in total CD4⁺ events (5 experiments; asterisk, *P* < 0.05; triple asterisk, *P* < 0.001; see representative cytometry plots in Supplementary Fig. 1). **d**, T–B conjugation assays were conducted using OT-2 cells transiently transfected with DNA constructs expressing either GFP or GFP-tagged SAP as indicated. Frequencies measured in 6 independent experiments are presented as mean ± s.e.m. (also see Supplementary Fig. 1).

Consistent with these latter results, we found that SAP(R78A), a mutant SAP capable of binding SLAM-related molecules but severely impaired in binding FYN kinase¹², rescued T–B cell conjugate formation *in vitro* (Supplementary Fig. 2). Taken together, our analyses of T–APC conjugation establish that SAP critically regulates the contact duration and the adhesive strength of cognate T–B cell but not antigen-driven T–DC interactions.

B cells do not acquire sufficient help

The cell numbers used in the preceding experiments can result in competition among T cells for antigen on presenting cells *in vivo*³², raising the question of whether *Sap*^{−/−} T cells would, if present alone, form more stable contacts with antigen-specific B cells or provide B cells with a summed contact time from sequential brief interactions comparable to the total interaction time with wild-type T cells. To examine these issues, small numbers (6×10^4) of *Sap*^{+/+} or *Sap*^{−/−} OT-2 T cells were separately transferred into recipient B6 mice that concomitantly received 3×10^5 MD4 B cells. Cumulative interactions between B and T cells were assessed by continuous intravital imaging over a 3-h period of time between 60 h and 72 h

post-immunization, when substantial clonal expansion has begun and physical co-localization of activated T and B cells can be observed^{8,29,33,34}. Under these modified conditions, *Sap*^{−/−} OT-2 cells still demonstrated abbreviated contacts with activated MD4 B cells after immunization with HEL–OVA (2.1 ± 0.2 min (mean \pm s.e.m.) versus 8.9 ± 0.8 min for *Sap*^{+/+} T cells, $P < 0.0001$; Fig. 3a and Supplementary Movie 4). Whereas many contacts involving *Sap*^{−/−} T cells were brief and similar to non-cognate T–B cell contacts in duration, on the population level they lasted for a statistically longer period of time (2.1 ± 0.2 versus 1.3 ± 0.1 min for non-cognate interactions, $P = 0.0014$), indicating that antigen-specific interactions occurred between *Sap*^{−/−} OT-2 cells and MD4 B cells but were not sustained. Individual B cells were also tracked for their interactions with different T cells. MD4 cells often accumulated >1 h of interactions with *Sap*^{+/+} OT-2 T cells, whereas they were rarely able to accumulate more than 20 min of contact with *Sap*^{−/−} OT-2 T cells over the same period (Fig. 3b–d and Supplementary Movies 4 and 5). B cells remained equally trackable under the two conditions and interacted with similar numbers of *Sap*^{−/−} or *Sap*^{+/+} T cells (Supplementary Fig. 3). The marked reduction in cumulative time

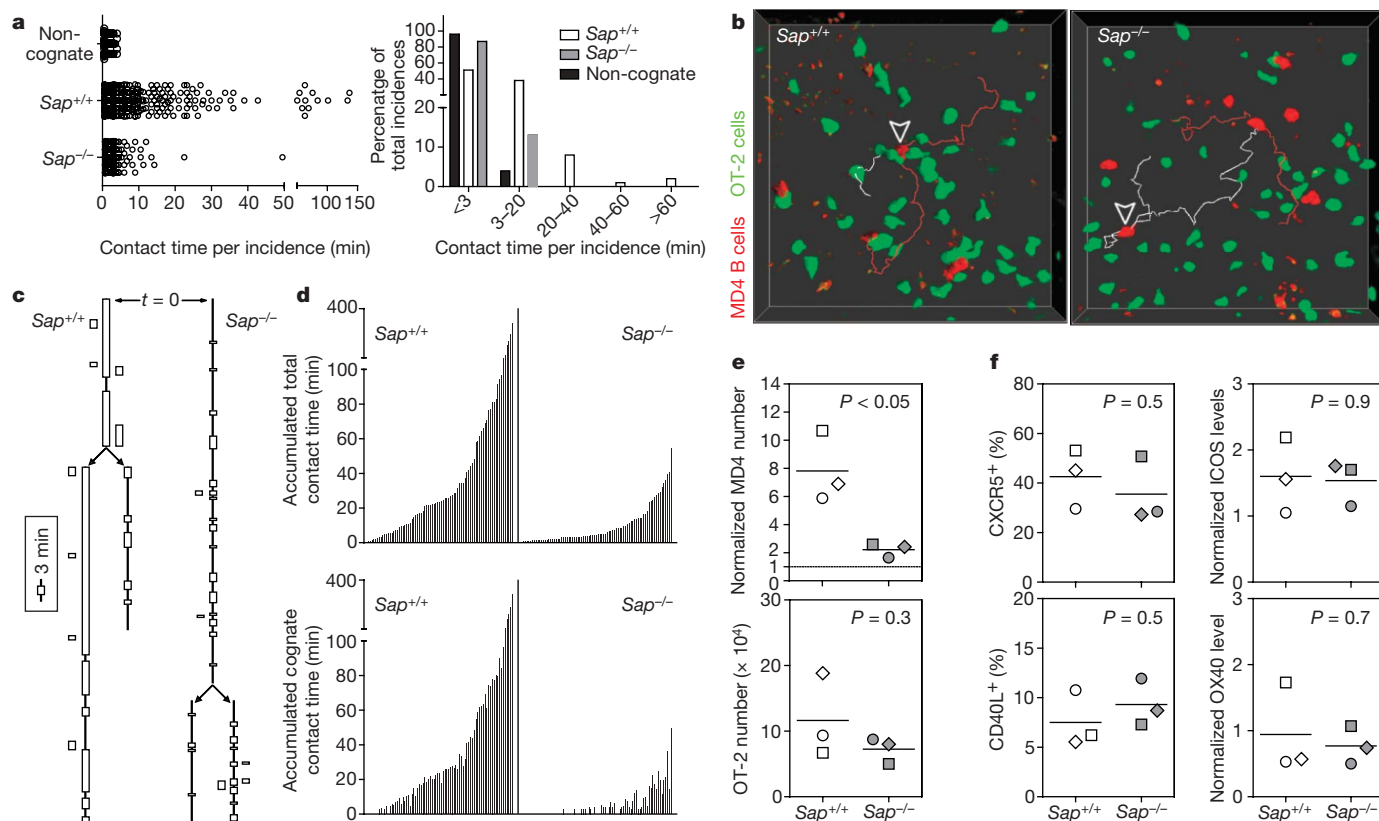


Figure 3 | B cells fail to receive contact-dependent help from SAP-deficient T cells. **a**, Individual T–B contact durations (left) and their distribution (right). Non-cognate conditions involve *Sap*^{+/+} OT-2 cells non-specifically interacting with MD4 B cells after immunization with HEL–BSA mixed with OVA protein (137 measurements from 2 experiments). Cognate conditions involve specific interactions between *Sap*^{+/+} or *Sap*^{−/−} OT-2 T cells and MD4 B cells after immunization with HEL–OVA (380 and 390 measurements from 3 experiments, respectively). **b–d**, Individual MD4 B cells tracked for cumulative interactions with OT-2 T cells over 3 h. **b**, Examples of tracked MD4 cells (arrowheads). MD4 cells represented by white tracks divided during the imaging period; daughter cells are represented by red tracks. **c**, Interaction histories of the MD4 cells in **b**. Vertical lines indicate the periods during which the B cells were present in the imaged volume. Arrows indicate cell divisions giving rise to daughter cells. Boxes represent one contact with a T cell; height indicates the contact duration (see also Supplementary Movie 5). Adjacent boxes indicate simultaneous interactions with more than one T cell. **d**, MD4 cells in each

condition tracked in 3 experiments (82 total). Each bar represents one MD4 cell. Bar heights are the sum of all T-cell contacts (top) or only those contacts longer than 3 min (bottom), a cut-off differentiating cognate from non-cognate T–B interactions with a 95% confidence as determined in **a**, **e**, **f**, MD4 B cells (7-amino-actinomycin D (7-AAD)[−]CD19⁺IgM⁺CD4[−]GFP[−] singlet events) and OT-2 T cells (7-AAD[−]CD19[−]CD4⁺GFP⁺ singlet events) analysed by FACS in 3 experiments at 96 h post HEL–OVA immunization. Individual symbols (open, *Sap*^{+/+}; filled, *Sap*^{−/−}), mean of 3–4 mice per group per experiment; lines, mean of the 3 experiments. **e**, Cell expansion. Top, numbers of MD4 cells recovered after co-transfer of *Sap*^{+/+} or *Sap*^{−/−} OT-2 T cells were normalized against the number obtained when no exogenous OT-2 cells were co-transferred (line at unit 1). Bottom, absolute numbers of recovered OT-2 cells after co-transfer with B cells. **f**, Surface expression of CXCR5, CD40L, ICOS and OX40 by the two types of OT-2 T cells. See Methods for details of quantitative analysis and Supplementary Fig. 4 for representative cytometry plots.

of contact with *Sap*^{-/-} OT-2 T cells correlated with a pronounced impairment of MD4 B cell clonal expansion (Fig. 3e). Nonetheless, *Sap*^{-/-} OT-2 T cells expanded comparably to their *Sap*^{+/+} counterparts and expressed similar levels of CD40L, ICOS, OX40 and CXCR5 (Fig. 3e, f and Supplementary Fig. 4) at this time. The latter is a phenotype considered characteristic of activated T cells capable of effectively promoting humoral responses and with features of T_{FH} precursor cells as commonly defined^{9,35}. In combination with the previous data, these findings suggest that it is not an intrinsic inability of SAP-deficient T cells to express key molecular signals, but rather a failure to exchange these signals during sufficiently long periods of contact with B cells, that accounts for their markedly reduced capacity to promote B-cell expansion and subsequent GC development.

Limited recruitment to and retention in GCs

Recruitment of helper T cells to a forming GC is necessary for maintaining the GC reaction³⁶, and GC-localized T_{FH} cells are required for effective immunoglobulin class switching and antibody affinity maturation⁹. To explore whether SAP deficiency also impairs T-cell recruitment to and retention within nascent GCs, we transferred equal numbers (3×10^4) of cyan fluorescent protein (CFP)-expressing *Sap*^{+/+} OT-2 cells and GFP-expressing *Sap*^{-/-} OT-2 cells into B6

mice together with 3×10^5 non-fluorescent MD4 cells. Six to eight days after HEL-OVA immunization, GCs developed within the draining lymph node, seen as GL7⁺ (LY77⁺) areas that largely excluded immunoglobulin D⁺ (IgD⁺) naive B cells (Supplementary Fig. 5). Whereas the IgD⁺ follicular mantle zone was populated by both *Sap*^{+/+} and *Sap*^{-/-} OT-2 cells, the GC area predominantly contained *Sap*^{+/+} OT-2 cells, consistent with a failure of GC recruitment and/or retention of *Sap*^{-/-} T cells.

To examine this issue dynamically, recipient mice were also given dye-labelled naive B cells 24 to 48 h before intravital imaging to differentiate follicular mantle and GC areas. Although individual naive B cells migrate in and out of GCs³⁷, as a population they remain substantially excluded from GCs (Supplementary Fig. 5), allowing simultaneous identification of the follicular area and approximation of the GC–mantle border using time-averaged images (see Supplementary Fig. 6 for details). *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells exhibited markedly different dynamic patterns within the follicle (Fig. 4a and Supplementary Movie 6). Whereas *Sap*^{+/+} OT-2 T cells migrated freely into and accumulated within the GCs, *Sap*^{-/-} OT-2 T cells mainly swarmed in the follicular mantle. To quantify migratory behaviours of the two types of T cells around GCs, a tessellation algorithm was used to reconstruct the GC surface (Supplementary

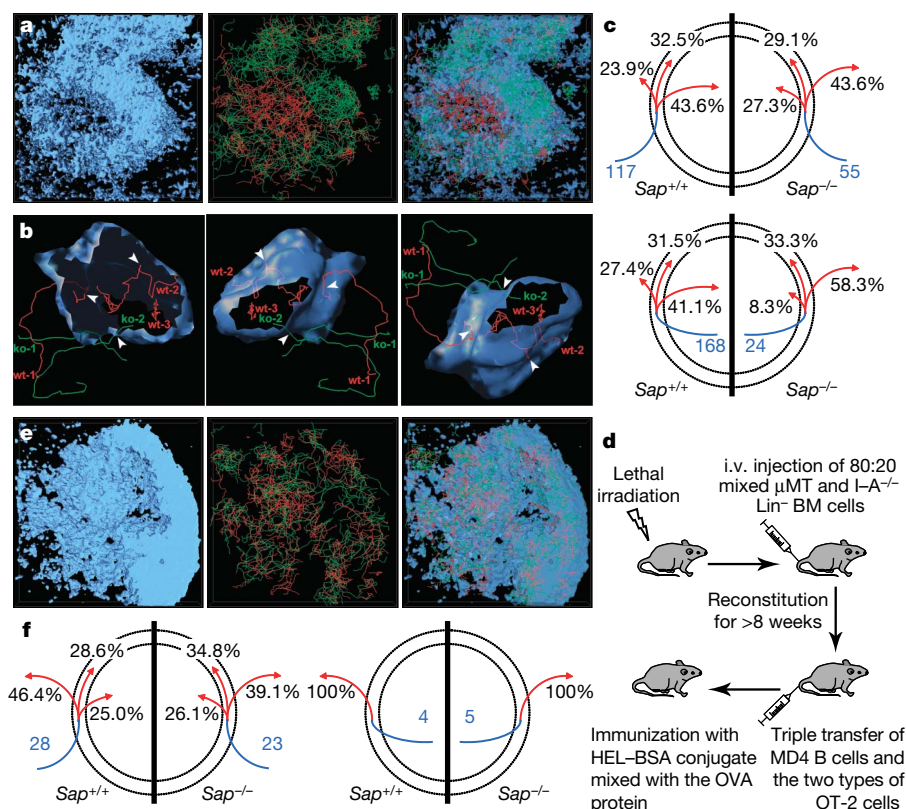


Figure 4 | Defective GC recruitment and retention of *Sap*^{-/-} T cells due to lack of efficient cognate interactions with B cells. **a**, Typical migration patterns of *Sap*^{+/+} (red) and *Sap*^{-/-} (green) OT-2 T cells in follicles containing cognate GCs. Left, the naive B-cell-dominated follicular mantle zone encasing GC area (see Supplementary Fig. 6 for details of three-dimensional rendering); middle, distribution of migration tracks of OT-2 T cells; right, overlay with the mantle zone rendered as semi-transparent (see also Supplementary Movie 6). **b**, GC surface identified from the data in **a** is rendered semi-transparent in three different orientations together with examples of five T-cell tracks differentially interacting with this surface. Tracks are labelled at their temporal beginnings. Arrowheads highlight positions where tracks cross the GC surface. ko, knockout; wt, wild type. **c**, Migration patterns of *Sap*^{+/+} and *Sap*^{-/-} T cells that arrived at a 10-μm-wide virtual mantle–GC interface zone. Area between the two dotted circles represents the 10-μm-wide interface zone; inner circle represents the GC surface, whereas the area outside of the outer circle is the follicular mantle.

Blue line segment indicates the region in which a track began (that is, the mantle zone (top) or the GC (bottom)). Numbers in blue are the total number of tracks of the indicated type analysed. Three red arrows denote tracks that subsequently moved into the mantle zone, stayed within the interface zone, or moved into the GC, respectively. Corresponding numbers denote percentages of tracks of each genotype that exhibited the indicated behaviour. Data are pooled from three experiments. **d–f**, Migration of *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells in follicles containing non-cognate GCs. **d**, A schematic showing the protocol used to generate GCs that are non-cognate to the OT-2 cells activated in the same lymph node. i.v., intravenous (see Methods for details). BM, bone marrow. **e**, Typical migration patterns of the two types of T cells in the same configuration as in **a** (see also Supplementary Movie 7). **f**, T-cell tracks that came within 10 μm of non-cognate GCs were analysed for their subsequent migrations using the same method as that used in **c**. Data from two experiments are pooled.

Fig. 6), and a virtual mantle–GC interface zone was then defined as encompassing spatial coordinates within 10 μm of the GC outer surface. T-cell tracks were subsequently classified according to their interactions with this interface zone (Fig. 4b, c). When reaching the interface from the follicular mantle, *Sap*^{+/+} OT-2 T cells were more likely to continue into the GC than to return to the mantle. *Sap*^{-/-} T cells exhibited the opposite behaviour, being more likely to turn back than to cross into the GC. Conversely, when reaching the interface zone from within the GC, *Sap*^{+/+} T cells were more likely to return to the GC than to escape into the mantle, whereas *Sap*^{-/-} T cells were much more likely to escape than to return. For those tracks that started from the interface zone, *Sap*^{-/-} T cells exhibited a preference for migrating into the mantle compared to moving into the GC, whereas for *Sap*^{+/+} T cells the movement in these two directions was comparable. Consistent with these distinct migration patterns, *Sap*^{-/-} T cells also exhibited significantly shorter GC retention times (Supplementary Fig. 7). Therefore, in the absence of SAP, T cells are not efficiently recruited into or retained within a nascent GC. These data suggest SAP-deficient T cells cannot act as effective GC-associated T_{FH} cells to sustain the GC reaction, a defect that would contribute to the profound impairment of GC responses in SAP-deficient hosts.

To address whether the inefficient GC recruitment and retention of *Sap*^{-/-} T cells resulted from their reduced antigen-specific interactions with B cells, radiation bone-marrow chimaeras were constructed so that endogenous B cells were deficient in MHC class II expression and thus unable to engage in antigen-specific interactions with T cells, whereas T cells could still interact with and be activated by non-B APCs including DCs (Supplementary Fig. 8; see a methodological diagram in Fig. 4d). After co-transfer of both *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells together with MD4 B cells into these chimaeric animals, they were immunized with a mixture of HEL–bovine serum albumin (BSA) and intact OVA proteins. In this setting, MD4 B cells cannot engage transferred OT-2 T cells as cognate partners but can still form GCs (Supplementary Fig. 9a) by using cognate help from endogenous CD4⁺ T cells. OT-2 cells in these mice can be normally activated by OVA-presenting DCs, but are deprived of cognate interactions with B cells, which are either transferred MD4 cells that do not present OVA or endogenous B cells that are derived from MHC-class-II-deficient bone marrow progenitors. Whereas both *Sap*^{+/+} and *Sap*^{-/-} OT-2 T cells upregulated CXCR5 and ICOS expression in this experimental system (Supplementary Fig. 9b), now they both failed to be efficiently recruited into and retained within GCs (Fig. 4e, f, Supplementary Fig. 7 and Supplementary Movies 7 and 8). When the same chimaeric mice were immunized with HEL–OVA, a condition in which OT-2 T cells could engage MD4 B cells as cognate partners, *Sap*^{+/+} OT-2 T cells were recruited to the GC, whereas *Sap*^{-/-} OT-2 T cells were not (data not shown). These findings indicate that T cells require cognate interactions with B cells to be recruited into and retained within GCs and suggest that, by failing to engage productively in such cell–cell interactions, SAP-deficient T cells are unable to physically localize to the GC to effectively sustain the GC reaction.

Conclusions

This study demonstrates that SAP expression in T cells is critical for stable antigen-dependent T–B cell adhesion but dispensable for T–DC interactions, revealing that distinct molecular rules govern the stable interactions of CD4⁺ T cells with two major APC types *in vivo*. Stabilization of T–B cell interactions may rely on a T-cell-autonomous mechanism orchestrated by SAP and its associated SLAM family molecules at the cell–cell interface, whereas stable T–DC association might be actively promoted by means of DC-mediated mechanisms, such as those involving cytoskeletal and membrane mechanics^{38,39}. In accordance with this notion, B cells but not DCs express high levels of SLAM (also known as CD150, SLAMF1), CD84, Ly9 (CD229) and Ly108 (SLAMF6) molecules

(Supplementary Fig. 10), which are also highly expressed by T cells activated *in vitro* and by T_{FH} cells *in vivo* (ref. 40 and data not shown). The differential SAP regulation of T–DC and T–B cell interactions correlates with our observations that CD4⁺ T-cell activation *in vivo* is grossly normal irrespective of SAP expression, whereas otherwise competent effector T cells fail to efficiently deliver contact-dependent help to cognate B cells in the absence of SAP. At the cellular level, these results provide a likely explanation for why SAP deficiency specifically impairs T-dependent humoral immunity but leaves intact or even exaggerates cell-mediated immune responses that also require CD4⁺ T-cell activity^{13–15,41,42}. Although additional factors may also contribute to the overall defects of humoral immunity seen in SAP-deficient hosts⁴³, our results strongly indicate that the predominant mechanism is the disruption of antigen-specific T–B adhesion. Future studies are necessary to determine the molecular basis for how SAP selectively controls cognate T–B association, to identify SAP-associated SLAM family proteins involved in this process, and to elucidate the cascade of intercellular signal exchange that transpires across a stable T–B synapse.

Our findings offer a compelling explanation for the profound impairment of the GC response in SAP-deficient mice and suggest a mechanism contributing to the B-cell-centric pathologies associated with the X-linked lymphoproliferative syndrome. They provide new insights into the operating principles of intercellular communication between T cells and diverse APCs and serve as a striking example of the importance of a direct visualization approach in understanding immune system function *in vivo*.

METHODS SUMMARY

Cell conjugation assay *in vitro*. OT-2 T cells (5×10^5 per well) were incubated for 30 min at 37 °C in 96-well U-bottom plates with DCs (10^6 per well) or lipopolysaccharide (LPS)-activated B cells (2×10^6 per well) that had been pulsed with antigen. Conjugates were enumerated by flow cytometry after the cell mixture was stained at 4 °C for CD4, CD11c (also known as ITGAX) and CD19 and repeatedly washed.

Adoptive transfer, cell phenotyping and intravital imaging. To visualize T–DC interactions (Fig. 1), 2×10^6 OVA₃₂₃-pulsed DCs per mouse were injected subcutaneously 24 h before intravenous transfer of *Sap*^{+/+} and *Sap*^{-/-} T cells (3×10^6 each). Imaging was conducted 12 to 24 h later. To examine activation phenotypes, 10^6 DCs and 2×10^5 GFP-expressing T cells per mouse were used. To visualize interactions between B cells and *Sap*^{+/+} and *Sap*^{-/-} T cells in the same lymph node (Fig. 2), 3×10^6 OT-2 T cells of each genotype were co-transferred into mice together with 5×10^6 wild-type B cells. Immunization was performed 12 h before cell transfer, and imaging was conducted 24 to 36 h thereafter. To visualize T–B interactions under non-competitive conditions and to assay T and B cell expansion (Fig. 3), 6×10^4 GFP-expressing OT-2 T cells were co-transferred together with 3×10^5 CFP-expressing B cells 24 h before immunization. Imaging and cytometric analyses were conducted 60–72 and 96 h later, respectively. To visualize GC recruitment and retention of T cells (Fig. 4), 3×10^4 CFP-expressing *Sap*^{+/+} and 3×10^4 GFP-expressing *Sap*^{-/-} OT-2 T cells were co-transferred with 3×10^5 non-fluorescent MD4 B cells. Imaging was conducted 6 to 8 days post immunization. Dye-labelled naive B cells ($2\text{--}4 \times 10^7$) were given 1 day before imaging to provide follicle and GC landmarks. The imaging set-up was essentially as described previously⁴⁴. For imaging sessions longer than 2 h, the animal's hydration was maintained by lactated Ringer's solution given by means of a catheter. The typical $x \times y \times z$ dimension was $0.5\text{--}1.1 \times 0.5\text{--}1.1 \times 3 \mu\text{m}$, and the time resolution was 30–45 s. For experiments involving co-transfer of two types of dye-labelled T cells, the cells were always reciprocally labelled to control for potential dye-induced behavioural differences.

Statistical analysis. The Mann–Whitney rank sum test was used to calculate *P* values for highly skewed distributions. For Gaussian-like distributions, two-tailed *t*-tests were used.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 July; accepted 15 August 2008.

- Coico, R. F., Bhogal, B. S. & Thorbecke, G. J. Relationship of germinal centers in lymphoid tissue to immunologic memory. VI. Transfer of B cell memory with

- lymph node cells fractionated according to their receptors for peanut agglutinin. *J. Immunol.* **131**, 2254–2257 (1983).
2. Berek, C., Berger, A. & Apel, M. Maturation of the immune response in germinal centers. *Cell* **67**, 1121–1129 (1991).
 3. Jacob, J., Kelsoe, G., Rajewsky, K. & Weiss, U. Intracloonal generation of antibody mutants in germinal centres. *Nature* **354**, 389–392 (1991).
 4. MacLennan, I. C. *et al.* The changing preference of T and B cells for partners as T-dependent antibody responses develop. *Immunol. Rev.* **156**, 53–66 (1997).
 5. Raff, M. C. Role of thymus-derived lymphocytes in the secondary humoral immune response in mice. *Nature* **226**, 1257–1258 (1970).
 6. Mitchison, N. A. The carrier effect in the secondary response to hapten–protein conjugates. II. Cellular cooperation. *Eur. J. Immunol.* **1**, 18–27 (1971).
 7. Lanzavecchia, A. Antigen-specific interaction between T and B cells. *Nature* **314**, 537–539 (1985).
 8. Garside, P. *et al.* Visualization of specific B and T lymphocyte interactions in the lymph node. *Science* **281**, 96–99 (1998).
 9. Vinuesa, C. G., Tangye, S. G., Moser, B. & Mackay, C. R. Follicular B helper T cells in antibody responses and autoimmunity. *Nature Rev. Immunol.* **5**, 853–865 (2005).
 10. Purtilo, D. T., Cassel, C. K., Yang, J. P. & Harper, R. X-linked recessive progressive combined variable immunodeficiency (Duncan's disease). *Lancet* **305**, 935–940 (1975).
 11. Grierson, H. L. *et al.* Immunoglobulin class and subclass deficiencies prior to Epstein–Barr virus infection in males with X-linked lymphoproliferative disease. *Am. J. Med. Genet.* **40**, 294–297 (1991).
 12. Ma, C. S., Nichols, K. E. & Tangye, S. G. Regulation of cellular and humoral immune responses by the SLAM and SAP families of molecules. *Annu. Rev. Immunol.* **25**, 337–379 (2007).
 13. Wu, C. *et al.* SAP controls T cell responses to virus and terminal differentiation of TH2 cells. *Nature Immunol.* **2**, 410–414 (2001).
 14. Czar, M. J. *et al.* Altered lymphocyte responses and cytokine production in mice deficient in the X-linked lymphoproliferative disease gene *SH2D1A/DSHP/SAP*. *Proc. Natl Acad. Sci. USA* **98**, 7449–7454 (2001).
 15. Crotty, S. *et al.* SAP is required for generating long-term humoral immunity. *Nature* **421**, 282–287 (2003).
 16. Morra, M. *et al.* Defective B cell responses in the absence of *SH2D1A*. *Proc. Natl Acad. Sci. USA* **102**, 4819–4823 (2005).
 17. Cannons, J. L. *et al.* SAP regulates T cell-mediated help for humoral immunity by a mechanism distinct from cytokine regulation. *J. Exp. Med.* **203**, 1551–1565 (2006).
 18. Coffey, A. J. *et al.* Host response to EBV infection in X-linked lymphoproliferative disease results from mutations in an SH2-domain encoding gene. *Nature Genet.* **20**, 129–135 (1998).
 19. Sayos, J. *et al.* The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM. *Nature* **395**, 462–469 (1998).
 20. Cannons, J. L. *et al.* SAP regulates T_H2 differentiation and PKC- θ -mediated activation of NF- κ B1. *Immunity* **21**, 693–706 (2004).
 21. Dent, A. L., Hu-Li, J., Paul, W. E. & Staudt, L. M. T helper type 2 inflammatory disease in the absence of interleukin 4 and transcription factor STAT6. *Proc. Natl Acad. Sci. USA* **95**, 13823–13828 (1998).
 22. Andoh, A. *et al.* Absence of interleukin-4 enhances germinal center reaction in secondary immune response. *Immunol. Lett.* **73**, 35–41 (2000).
 23. Huang, A. Y., Qi, H. & Germain, R. N. Illuminating the landscape of *in vivo* immunity: insights from dynamic *in situ* imaging of secondary lymphoid tissues. *Immunity* **21**, 331–339 (2004).
 24. Stoll, S., Delon, J., Broetz, T. M. & Germain, R. N. Dynamic imaging of T cell–dendritic cell interactions in lymph nodes. *Science* **296**, 1873–1876 (2002).
 25. Mempel, T. R., Henrickson, S. E. & Von Andrian, U. H. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* **427**, 154–159 (2004).
 26. Miller, M. J., Safrina, O., Parker, I. & Cahalan, M. D. Imaging the single cell dynamics of CD4⁺ T cell activation by dendritic cells in lymph nodes. *J. Exp. Med.* **200**, 847–856 (2004).
 27. Celli, S., Lemaitre, F. & Bousso, P. Real-time manipulation of T cell–dendritic cell interactions *in vivo* reveals the importance of prolonged contacts for CD4⁺ T cell activation. *Immunity* **27**, 625–634 (2007).
 28. Abbas, A. K., Haber, S. & Rock, K. L. Antigen presentation by hapten-specific B lymphocytes. II. Specificity and properties of antigen-presenting B lymphocytes, and function of immunoglobulin receptors. *J. Immunol.* **135**, 1661–1667 (1985).
 29. Okada, T. *et al.* Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS Biol.* **3**, e150 (2005).
 30. Davidson, D. *et al.* Genetic evidence linking SAP, the X-linked lymphoproliferative gene product, to Src-related kinase FynT in T_H2 cytokine regulation. *Immunity* **21**, 707–717 (2004).
 31. McCausland, M. M. *et al.* SAP regulation of follicular helper CD4 T cell development and humoral immunity is independent of SLAM and Fyn kinase. *J. Immunol.* **178**, 817–828 (2007).
 32. Garcia, Z. *et al.* Competition for antigen determines the stability of T cell–dendritic cell interactions during clonal expansion. *Proc. Natl Acad. Sci. USA* **104**, 4553–4558 (2007).
 33. Kelsoe, G. & Zheng, B. Sites of B-cell activation *in vivo*. *Curr. Opin. Immunol.* **5**, 418–422 (1993).
 34. Liu, Y. J. *et al.* Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. *Eur. J. Immunol.* **21**, 2951–2962 (1991).
 35. Moser, B., Schaerli, P. & Loetscher, P. CXCR5⁺ T cells: follicular homing takes center stage in T-helper-cell responses. *Trends Immunol.* **23**, 250–254 (2002).
 36. de Vinuesa, C. G. *et al.* Germinal centers without T cells. *J. Exp. Med.* **191**, 485–494 (2000).
 37. Schwickert, T. A. *et al.* *In vivo* imaging of germinal centres reveals a dynamic open structure. *Nature* **446**, 83–87 (2007).
 38. Al-Alwan, M. M., Rowden, G., Lee, T. D. & West, K. A. The dendritic cell cytoskeleton is critical for the formation of the immunological synapse. *J. Immunol.* **166**, 1452–1456 (2001).
 39. Benvenuti, F. *et al.* Requirement of Rac1 and Rac2 expression by mature dendritic cells for T cell priming. *Science* **305**, 1150–1153 (2004).
 40. Chtanova, T. *et al.* T follicular helper cells express a distinctive transcriptional profile, reflecting their role as non-Th1/Th2 effector cells that provide help for B cells. *J. Immunol.* **173**, 68–78 (2004).
 41. Hron, J. D. *et al.* *SH2D1A* regulates T-dependent humoral autoimmunity. *J. Exp. Med.* **200**, 261–266 (2004).
 42. Crotty, S. *et al.* Hypogammaglobulinemia and exacerbated CD8 T-cell-mediated immunopathology in SAP-deficient mice with chronic LCMV infection mimics human XLP disease. *Blood* **108**, 3085–3093 (2006).
 43. Ma, C. S. *et al.* Impaired humoral immunity in X-linked lymphoproliferative disease is associated with defective IL-10 production by CD4⁺ T cells. *J. Clin. Invest.* **115**, 1049–1059 (2005).
 44. Qi, H., Egen, J. G., Huang, A. Y. & Germain, R. N. Extrafollicular activation of lymph node B cells by antigen-bearing dendritic cells. *Science* **312**, 1672–1676 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements H.Q. is in debt to Y. Hong for support, encouragement and inspiration. This work was funded by the intramural research programs of the National Institute of Allergy and Infectious Disease and National Human Genome Research Institute, National Institutes of Health, USA.

Author Contributions H.Q. and J.C. conducted the experiments. R.N.G. and P.L.S. contributed equally to this study. All authors contributed collectively to designing the experiments, interpreting the data and writing the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.N.G. (rgermain@niaid.nih.gov) or P.L.S. (pams@nhgri.nih.gov).

METHODS

Mice. SAP-deficient mice were described previously¹⁴ and backcrossed to B6 for at least ten generations. B6 (Jax 664), CD45.1 congenic (Jax 2014), HEL-specific Ig-transgenic MD4 (Jax 2595), μ MT (Jax 2288), GFP-expressing (Jax 4353), CFP-expressing (Jax 4218) and OVA_{323–339}-specific TCR-transgenic OT-2 (Jax 4194) mice were purchased from the Jackson Laboratory. I-A β ^{−/−} mice were purchased from Taconic Farms. Relevant mice were interbred to obtain CFP-MD4, *Sap*^{+/+} CFP-OT-2, and *Sap*^{+/+} or *Sap*^{−/−} GF-OT-2 mice. All mice were maintained under specific-pathogen free conditions, and used in accordance of NIH institutional guidelines for animal welfare.

Cell preparation, antigen generation and immunization. To isolate DCs, mouse spleens were digested with 400 μ g ml^{−1} Liberase CI and 20 μ g ml^{−1} DNase I for 30 min (Roche) before subjected to CD11c microbeads-based isolation protocol (Miltenyi Biotec). DCs were pulsed with indicated doses of OVA₃₂₃ peptide antigen in the presence of LPS for 2 h, washed, and used for the *in vitro* cell conjugation assay or for subcutaneous injection into mice after being labelled with 75 μ M Cell Tracker Blue (CMF₂HC). OT-2 T cells and B cells of either a MD4 or a polyclonal repertoire were isolated by CD4 T-cell isolation kit and naive B-cell isolation kit (Miltenyi Biotec), respectively. For intravital imaging experiments involving dye-labelled cells, the MD4 B cells were always labelled with 75 μ M CMF₂HC, whereas *Sap*^{+/+} and *Sap*^{−/−} OT-2 T cells were interchangeably labelled with either 2 μ M carboxyfluorescein diacetate succinimidyl ester (CFSE) or 4 μ M Cell Tracker Red (CMTPX). All dyes were purchased from Invitrogen. Model antigens used for subcutaneous immunization included OVA protein and chemical conjugates of OVA and HEL or BSA and HEL. The purified proteins were purchased from Sigma-Aldrich, and conjugates were made with the HydraLink heterobifunctional conjugation kit (SoluLink) according to manufacturer's instructions. In brief, HEL was modified by 6-hydrazinonicotinamide and then allowed to react at a molar ratio of 5:1 with OVA or BSA that was modified by 4-formylbenzoate. Excessive unconjugated HEL was then removed by size-exclusion chromatography. Mice were immunized with the indicated antigen mixed with alum and 0.2 μ g LPS.

Transient transfection of T cells. For certain experiments, blasting OT-2 T cells were transfected with 4 μ g DNA constructs expressing GFP-tagged SAP or GFP alone by an Amaxa nucleofector (program X-001) before being analysed for conjugation formation with B cells. Reconstituted expression of SAP proteins in *Sap*^{−/−} T cells was verified by western blotting.

Construction of mixed bone marrow chimaeras. B6 or CD45.1 congenic mice were lethally irradiated by γ -ray from a cesium source (900 rad) and then received intravenous transfer of a total of 1.25×10^5 bone-marrow-lineage-negative cells, which were isolated using the mouse lineage depletion kit (Miltenyi Biotec). To make chimaeric mice that lacked MHC class II, a mixture of 80% μ MT and 20% I-A β ^{−/−} donor bone marrow cells were used. For certain control experiments, I-A β ^{−/−} donor cells were replaced with regular B6 cells.

Flow cytometry and immunohistochemistry. All reagents were purchased from BD Pharmingen unless indicated otherwise. To examine surface phenotypes of OT-2 T cells or MD4 B cells, draining lymph node cells were washed, incubated with 50 μ g ml^{−1} rat and hamster IgG whole molecules (Pierce) and 10% 2.4G2

hybridoma supernatants, and then stained with indicated antibodies in PBS supplemented with 2 mM EDTA and 0.5% fetal calf serum. Staining reagents included AlexaFluor 700 anti-CD4 (eBioscience), AlexaFluor 700 CD38 (eBioscience), APC-Cy7 anti-CD19, phycoerythrin (PE) anti-Fas, PE anti-ICOS, PE anti-OX40, PE anti-I-A β , APC anti-CD11c, biotinylated anti-CD40L, biotinylated anti-CXCR5, biotinylated anti-IgM^a, 7-AAD, streptavidin APC, and isotype-matched non-specific antibodies. Cells were stained on ice with primary reagents for 1 h and with the secondary reagent for 20 min. Cytometric data were collected on an LSR II cytometer (BD) and analysed with FlowJo software (TreeStar). Dying cells and cell multiplets were excluded from analysis based on 7-AAD fluorescence and forward-scattering intensity (FSC)-height/FSC-area. Quantitative analysis of cell surface phenotypes involved determining percentages of positive cells based on a cutoff established by the isotype control (for CXCR5 and CD40L) or by calculating the normalized expression level based on geometric mean fluorescence intensities ($\log_{10}(\text{MFI}^{\text{sample}})/\log_{10}(\text{MFI}^{\text{isotype}})$ for ICOS and OX40). For immunohistochemical staining of lymph node sections, the protocol was as described previously¹⁴. Staining reagents included AlexaFluor 647 anti-B220, biotinylated anti-IgD, purified GL7 (eBioscience), AlexaFluor 647 anti-rat IgM (Invitrogen) and streptavidin AlexaFluor 568 (Invitrogen). All stained slides were mounted with ProlongGold antifade reagents (Invitrogen) and examined with a Leica TCS SP5 confocal system.

Intravital imaging and quantitative data analysis. To maintain consistency of visual presentation in all figures and movies, the *Sap*^{+/+} and *Sap*^{−/−} T cells were pseudo-coloured in red and green, respectively. An excitation wavelength between 783 and 800 nm was used for dye-labelled cells. For imaging experiments involving GFP- or CFP-expressing transgenic T or B cells, with or without additional CMTPX-labelled naive B cells, an excitation wavelength between 840 and 850 nm was used, a compromise wavelength for simultaneous visualization of CFP and GFP or all of the three fluorochromes. Whereas the overlap between normalized emission curves of CFP and GFP is substantial, the GFP transgenic line used here (Jax 4353, Jackson Laboratory) is ~100 times 'brighter' than the CFP transgenic line (Jax 4218, Jackson Laboratory), allowing real GFP signals to be readily distinguished from bleed-through signals from CFP. Post data acquisition, four-dimensional image data sets were analysed using Bitplane Imaris software package. Cell-cell contacts were scored manually in blind analyses. Cell migration was analysed through automatic cell tracking aided by manual supervision. Only cell tracks that lasted longer than 5 min were included in quantitative analysis. For certain experiments, it was necessary to quantitatively analyse cell positioning over time in relationship to the GC in space. To achieve this, the boundary of GC was manually traced for each optical slice in a time-averaged series and then mathematically defined as a triangulated surface in three-dimensional space by an optimized Dirichlet tessellation algorithm implemented in C++ (F.K. *et al.*, manuscript in preparation). The distance from a cell at any give time to the GC surface can then be calculated (see Supplementary Fig. 6 for a flow chart of the relevant data processing). For all imaging experiments, final presentations of time-lapse image sequences as stills or movies were created using Photoshop, AfterEffect and Illustrator (Adobe).

ARTICLES

Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing

Eleni P. Mimitou¹ & Lorraine S. Symington¹

DNA ends exposed after introduction of double-strand breaks (DSBs) undergo 5'–3' nucleolytic degradation to generate single-stranded DNA, the substrate for binding by the Rad51 protein to initiate homologous recombination. This process is poorly understood in eukaryotes, but several factors have been implicated, including the Mre11 complex (Mre11–Rad50–Xrs2/NBS1), Sae2/CtIP/Ctp1 and Exo1. Here we demonstrate that yeast Exo1 nuclease and Sgs1 helicase function in alternative pathways for DSB processing. Novel, partially resected intermediates accumulate in a double mutant lacking Exo1 and Sgs1, which are poor substrates for homologous recombination. The early processing step that generates partly resected intermediates is dependent on Sae2. When Sae2 is absent, in addition to Exo1 and Sgs1, unprocessed DSBs accumulate and homology-dependent repair fails. These results suggest a two-step mechanism for DSB processing during homologous recombination. First, the Mre11 complex and Sae2 remove a small oligonucleotide(s) from the DNA ends to form an early intermediate. Second, Exo1 and/or Sgs1 rapidly process this intermediate to generate extensive tracts of single-stranded DNA that serve as substrate for Rad51.

DSBs are potentially lethal lesions that can occur spontaneously during normal cell metabolism or by treatment of cells with DNA-damaging agents. If unrepaired or repaired inappropriately, DSBs can lead to mutagenic events such as chromosome loss, deletions, duplications or translocations. The Mre11 complex (Mre11–Rad50–Xrs2/NBS1) is rapidly recruited to DSBs, signals checkpoint activation through Tel1/ATM kinase and regulates 5'–3' resection of the DNA ends^{1–3}. The intrinsic Mre11 nuclease activity is essential for processing ends blocked by covalent adducts, such as Spo11 or hairpin-capped ends, but appears to be less important for processing HO endonuclease-induced DSBs^{4–7}. Sae2/CtIP/Ctp1 interacts with the Mre11 complex, and Sae2 deficiency is phenotypically similar to loss of the Mre11 nuclease activity^{6–10}. The functions of the Mre11 complex and Sae2 in processing DSBs are partly redundant with the 5'–3' double-strand exonuclease, Exo1 (ref. 11). In the absence of Exo1 and Sae2, 5'–3' resection of DSBs is reduced, but resection and homologous recombination still occur, suggesting the existence of at least one other factor in DSB processing¹². Here we demonstrate a function for the *Saccharomyces cerevisiae* RecQ helicase homologue, Sgs1, in DSB processing. As the Sgs1 homologue in humans, BLM, maintains genome integrity and guards against cancer predisposition¹³, our findings suggest that some of the defects observed in individuals with Bloom's syndrome could be due to altered processing of DSBs.

Sae2 and Exo1 function at different steps in resection

We developed a single-strand annealing assay (SSA), which is inherently sensitive to defects in DSB resection, to identify genes required for this process¹⁴. The assay uses tandem repeats of the *ade2* gene, one of which contains a recognition site for the I-SceI endonuclease (Fig. 1a)¹⁵. The strain also harbours a *GAL1p-I-SCEI* fusion to provide regulated expression of the nuclease. After induction of I-SceI, homology-dependent repair can occur by gene conversion directed by the uncut allele, or by SSA⁴. SSA requires extensive 5'–3' degradation (about 7 kilobases (kb) in this case) to expose complementary single-stranded DNA (ssDNA) followed by Rad52-dependent

annealing of the repeats. After trimming residual ssDNA tails and gap filling, recombinants form with loss of one of the repeats and the intervening sequence. In the *rad51Δ* mutant, gene conversion is eliminated, leaving repair of the DSB by SSA as the alternative mode¹⁶. The efficiency of SSA can be assessed by the plating efficiency of strains on medium containing galactose (YP_{Raf}/Gal) to induce constitutive expression of I-SceI. Changes in the kinetics of SSA can be evaluated by Southern blot analysis of NheI/EagI-digested genomic DNA.

The plating efficiency of the *rad51Δ* strain on YP_{Raf}/Gal was the same as on medium containing glucose (YPD), and Ade⁺ recombinants were evident by formation of white colonies (Fig. 1b). By physical analysis, DSB fragments appeared 1 h after inducing I-SceI expression and were then processed to form the SSA product visible at 2 h and accumulating to 80% of the total DNA after 24 h (Fig. 1c). 5'–3' resection past the EagI and NheI sites results in loss of the cut fragments. Therefore, increased stability of the cut fragments indicates a delay or reduced rate of processing. Although the DSB fragments produced in the *rad51Δ exo1Δ* mutant were processed with similar kinetics as in the *rad51Δ* mutant, there was a delay in SSA product formation and reduced yield. In the *rad51Δ sae2Δ* strain, the DSB fragments persisted for longer, but the deletion products still appeared at 2 h and accumulated to 50% of the total. The phenotypes of *sae2Δ* and *exo1Δ* mutants, stability of cut fragments and delay in product formation, respectively, were additive in the *rad51Δ exo1Δ sae2Δ* mutant strain, resulting in a substantial defect in SSA and a twofold reduction in plating efficiency on YP_{Raf}/Gal plates (Fig. 1). To characterize the formation of ssDNA, DNA samples were applied to nylon membranes in the native or denatured state, and tested for ability to hybridize to a probe adjacent to the DSB. This analysis revealed similar amounts of ssDNA adjacent to the break site in *rad51Δ exo1Δ* and *rad51Δ* strains at early time points, whereas a decrease was observed in the *rad51Δ sae2Δ* mutant (Fig. 1d). ssDNA persisted in the *rad51Δ exo1Δ* mutant, consistent with the observed delay in SSA product formation. These results are in line with a model in which Sae2 starts DSB processing to produce an early

¹Department of Microbiology, Columbia University Medical Center, 701 West 168th Street, New York, New York 10032, USA.

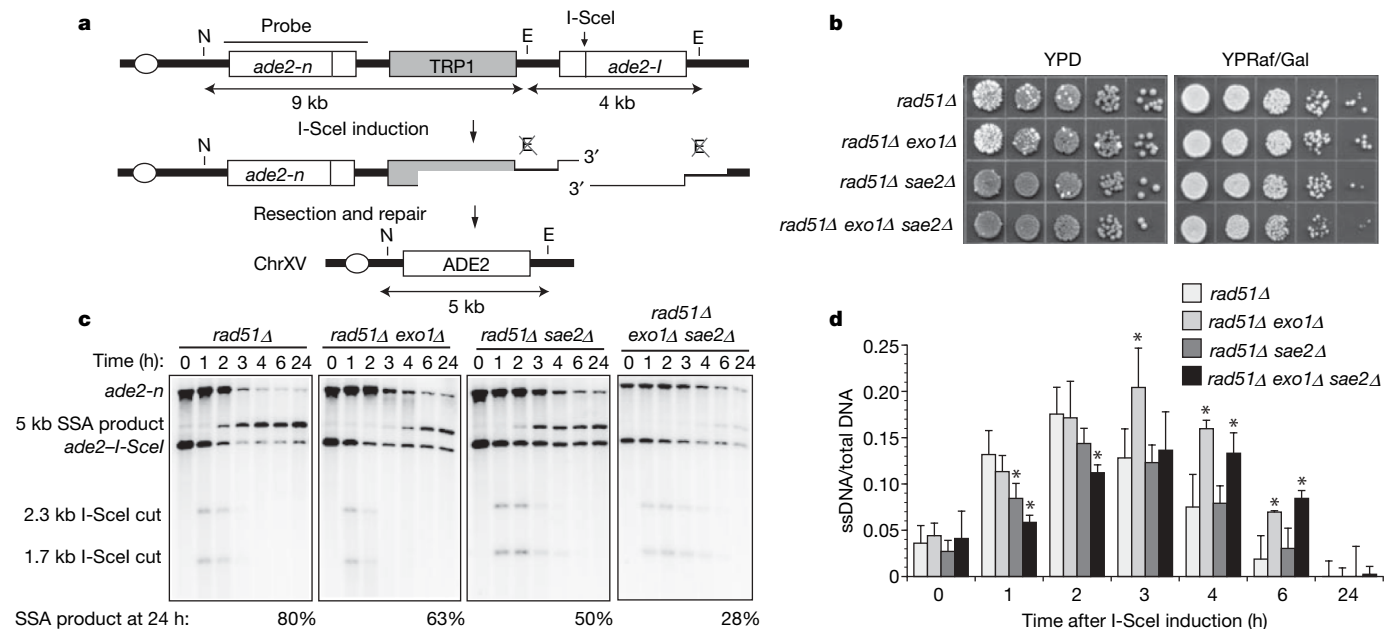


Figure 1 | Sae2 and Exo1 function at different steps in DSB resection.

a, Map of the chromosome XV region containing the *ade2* direct repeat. The *ade2-I* allele is cleaved by I-SceI to yield the 2.3- and 1.7-kb cut fragments that disappear over time, concomitant with repair to the 5-kb SSA product. **b**, SSA efficiency as indicated by sectoring on YPD plates and survival on YPRaf/Gal plates. Ade⁺ (white) colonies arise from *ade2* (red) colonies by SSA during growth on YPD owing to leaky expression of I-SceI. **c**, SSA

physical assay: monitoring of resection and repair of the I-SceI-induced DSB by Southern blot analysis of NheI/EagI-digested genomic DNA.

d, Quantification of the ADE2 hybridization signal on native versus denatured DNA samples bound on nylon membrane. The means from multiple experiments from three independent inductions are presented, the error bars indicate standard deviations; **P* < 0.02 (unpaired *t*-test).

intermediate that is rapidly matured by the processive nucleolytic activity of Exo1. In the absence of Sae2, a smaller subset of DSBs is acted on directly by Exo1 yielding SSA products at the same time, but with lower yield than found in the presence of Sae2.

DSB processing requires Sgs1 in the absence of Exo1

The presence of SSA products in the *exo1Δ sae2Δ* mutant suggests at least one other activity is operative in DSB processing. The missing

factor is not Mre11 because SSA products still form in *mre11Δ*, *mre11Δ exo1Δ*, and *mre11Δ exo1Δ sae2Δ* mutants (Supplementary Fig. 1). The nuclease-defective *mre11-H125N* mutant was slightly less defective than *sae2Δ*, by itself or with *exo1Δ* (data not shown). One possible mechanism for DSB processing would involve a helicase in conjunction with a single-strand-specific endo- or exonuclease. In *Escherichia coli*, the RecQ 3'–5' helicase and RecJ 5'–3' exonuclease function in DSB resection in the absence of the dominant RecBCD

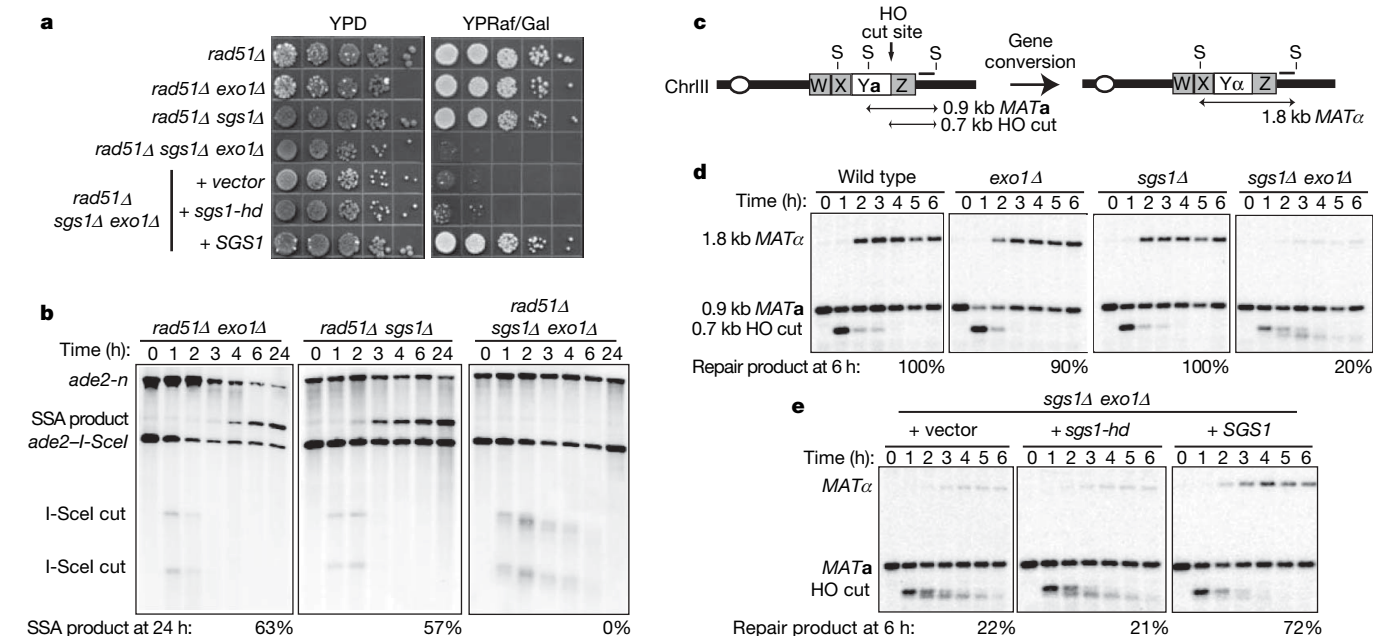


Figure 2 | Sgs1 and Exo1 function redundantly in DSB resection. **a**, SSA plating efficiency as described in Fig. 1b. **b**, SSA physical assay as described in Fig. 1c. **c**, Representation of MAT switching: the 0.9-kb MATa StyI fragment is cleaved by HO, resulting in a 0.7-kb cut fragment that disappears over time

concomitant with repair to a 1.8-kb MATα fragment. **d**, MAT switching assay: HO cut fragment processing and gene conversion efficiency as monitored by Southern blot analysis of StyI-digested genomic DNA. **e**, MAT switching assay on *sgs1Δ exo1Δ* cultures complemented with mutant or wild-type SGS1.

processing activity¹⁷. Therefore, we considered the possibility that the RecQ homologue of *S. cerevisiae*, Sgs1 (refs 18, 19), might fulfil a similar redundant function. The *rad51Δ sgs1Δ* mutant showed a delay in SSA (Fig. 2 and Supplementary Fig. 2); however, the *rad51Δ exo1Δ sgs1Δ* strain was completely deficient, as evidenced by the greater than 1,000-fold decrease in plating efficiency on YPRaf/Gal medium (Fig. 2a). Furthermore, there was no detectable deletion product by Southern blot analysis, even 24 h after I-SceI induction (Fig. 2b). Notably, in the *rad51Δ exo1Δ sgs1Δ* mutant novel intermediates were observed with slightly faster mobility than the DSB cut fragments. These became more smeared over time, suggesting very slow processing. We were unable to analyse the *sae2Δ sgs1Δ* double mutant because of a synthetic growth defect²⁰.

The SSA assay requires more extensive resection than is likely to be required for Rad51-dependent strand invasion. To determine whether the partly processed intermediates that accumulate in *exo1Δ sgs1Δ* strains are proficient for gene conversion, we analysed DSB processing in *RAD51* derivatives by a mating-type (*MAT*) switching assay. In this assay, DNA samples are digested with StyI to detect HO endonuclease cleavage and gene conversion products by Southern blot analysis (Fig. 2c)²¹. The *exo1Δ* and *sae2Δ* mutants exhibited slight defects in the efficiency of repair, whereas the *sgs1Δ* single mutant appeared normal (Fig. 2d and Supplementary Fig. 3). In the *exo1Δ sgs1Δ* double mutant, the DSB fragment was processed to yield an intermediate with faster mobility than the cut fragment. This persisted for several hours, as we had observed in the SSA assay. Despite the defect in DSB processing, the 1.8-kb gene-conversion product was still detected, although with reduced efficiency (20%) compared with wild type and the single mutants. This suggests very short ssDNA tails are sufficient to initiate Rad51-dependent gene conversion. To determine whether the helicase activity of Sgs1 is important for DSB processing, plasmids expressing wild-type or helicase-defective alleles of *SGS1* were used to transform the *exo1Δ sgs1Δ* double mutant, and the resulting strains tested for *MAT* switching or SSA proficiency²². As anticipated, the *SGS1* plasmid rescued the SSA and gene conversion defects of the *sgs1Δ exo1Δ* mutant, but the helicase-defective *sgs1-K706A* allele failed to complement *sgs1Δ* in these assays (Fig. 2). Previous studies have shown the steady-state level of the Sgs1-K706A protein to be similar to wild type²². Therefore, the helicase activity is required for DSB processing by Sgs1. The lack of full complementation by *SGS1* in the *MAT*-switching assay could be due to plasmid loss in some of the cells in the population.

Resection is attenuated in the *exo1Δ sgs1Δ* strain

To measure resection of 5' strands directly, we analysed HO-induced DSB fragments by restriction digestion and alkaline gel electrophoresis (Fig. 3a). Processing of the HO-induced break renders the DNA single-stranded and resistant to digestion with restriction enzymes, resulting in a ladder of bands of higher molecular mass²¹. By using RNA probes specific for the 3' and 5' strands to the right of the HO cut site, resection of the strands can be monitored. A *rad51Δ* background was used for this analysis to prevent engagement of the intermediates in repair.

The probe specific for 3' strands detected extensive resection of the HO-induced break in the *rad51Δ* strain (Fig. 3b). In the *rad51Δ exo1Δ* mutant, the first two ssDNA intermediates were formed with normal kinetics. However, the products indicative of extensive resection were delayed, resulting in accumulation of the r2 intermediate, suggesting Sgs1-dependent resection is less processive (Supplementary Fig. 4)⁵. The extent and timing of resection in the *rad51Δ sgs1Δ* mutant were similar to the *rad51Δ* strain. In contrast, only a faint band corresponding to resection beyond the first StyI site was detected in the *rad51Δ exo1Δ sgs1Δ* strain. Detection of a discrete band corresponding to degradation beyond the first StyI site and failure to detect products smaller than the cut fragment suggest the 3' ends are intact in the *rad51Δ exo1Δ sgs1Δ* mutant. This was confirmed by using a probe specific for the 5' strand. No intermediates

were detected in the *rad51Δ exo1Δ* or *rad51Δ sgs1Δ* mutants, presumably because of the rapid degradation of 5' ends. However, a smeared product was detected with faster migration than the 700-base-pair (bp) cut fragment in the *rad51Δ exo1Δ sgs1Δ* mutant, consistent with the severe resection defect. Interestingly, the smear does not occur immediately below the cut fragment, suggesting there is an initial endonuclease cleavage step that removes about 50–100 nucleotides from the 5' end.

Two-step mechanism for DSB processing

The results presented above suggest that after formation of a DSB, the ends are trimmed by an endonuclease to an intermediate form, which is subsequently processed by Exo1 or Sgs1. The Mre11 complex and Sae2 are likely to be involved in the initial processing step because DSB fragments are stabilized in *mre11Δ* and *sae2Δ* mutants, and recent studies have shown Mre11-dependent processing of DSBs to form short ssDNA oligonucleotides²³. Furthermore, Sae2 has an endonuclease activity that is stimulated by the Mre11 complex²⁴. We attempted to generate *exo1Δ sgs1Δ mre11Δ* or *exo1Δ sgs1Δ sae2Δ* triple mutants, but these were inviable (Supplementary Fig. 5a, b); the *exo1Δ sgs1Δ mre11-H125N* triple mutant was also inviable. A conditional *sae2* allele was created in the *exo1Δ sgs1Δ* background by expressing endogenous *SAE2* from the *P_{tet}* promoter,

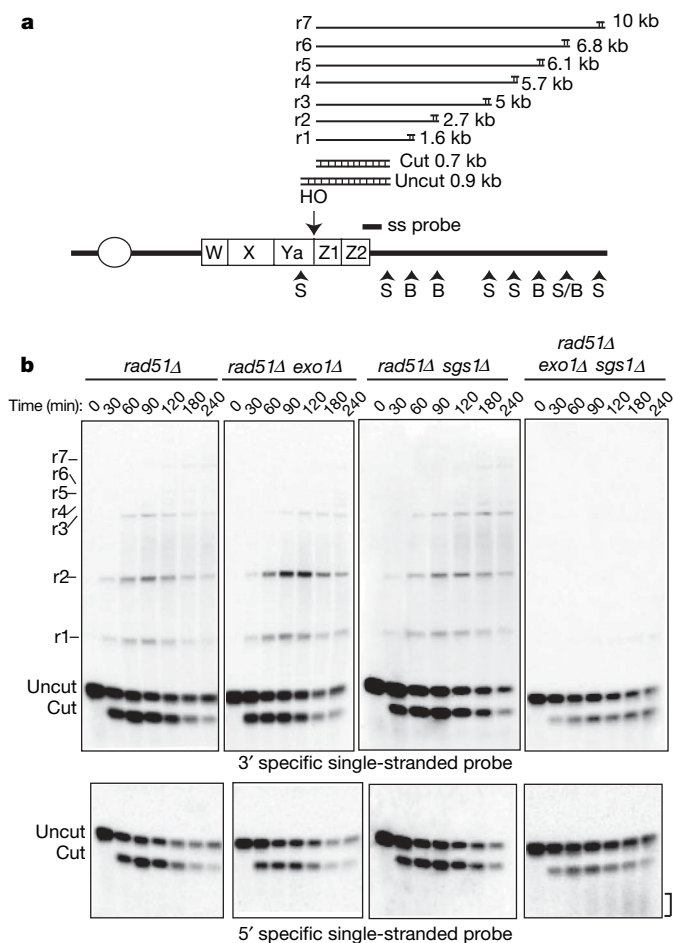


Figure 3 | Single-stranded intermediates fail to form in the absence of Exo1 and Sgs1. **a**, Representation of the method used to detect single-stranded intermediates after 5'–3' resection of the *MAT* locus, showing the positions of the HO cut site, the probe and the StyI (S) and BstXI (B) sites. Processing of the HO-induced break renders the DNA single-stranded and resistant to digestion with restriction enzymes, giving rise to distinct sets of fragments²¹. **b**, Alkaline electrophoresis of StyI/BstXI-digested genomic DNA: ssDNA intermediates formed after the resection of the HO-induced break were detected by using 3'- or 5'-specific riboprobes. The bracket indicates the smeared 5'-terminated strands observed in the *rad51Δ exo1Δ sgs1Δ* mutant.

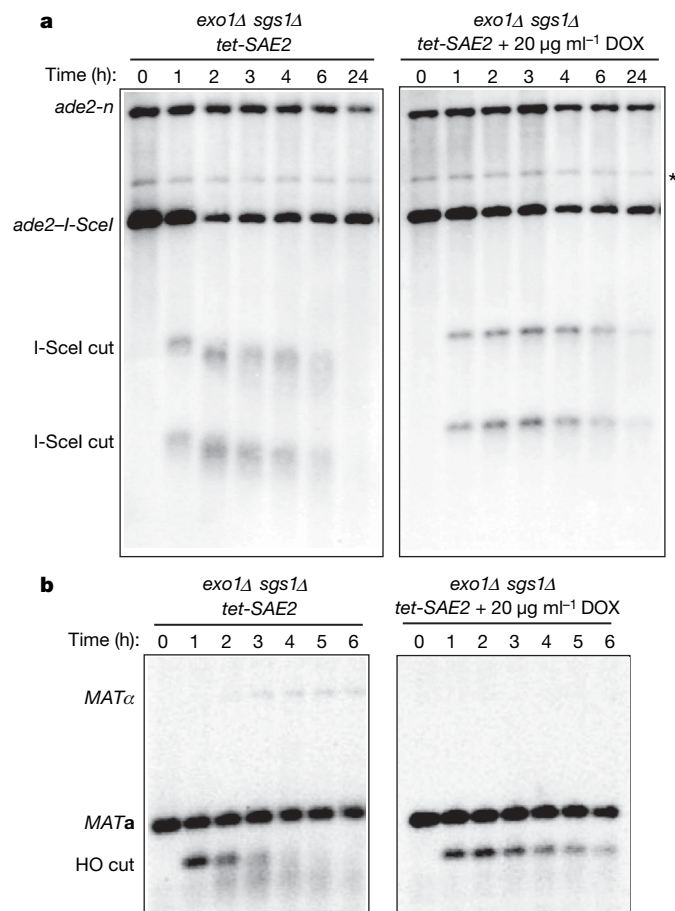


Figure 4 | Sae2 is required for creating the minimally resected intermediates. SSA assay (a) and mating-type switching assay (b) in cells treated or untreated with doxycycline before I-SceI or HO expression. The asterisk refers to a cross-hybridizing band, which is present at time 0 h; if it were due to recombinants in the population, it would allow survival of cells on YPRaf/Gal medium.

this strain also expresses the tetracycline-controlled trans-activator (tTA) to repress transcription of *SAE2* by growth in the presence of doxycycline²⁵. The *exo1Δ sgs1Δ P_{tet}-SAE2* strain was viable in the absence of doxycycline, indicating normal function. By western blot analysis, the Sae2 protein was depleted 4 h after addition of doxycycline to the growth medium (Supplementary Fig. 5c). In the absence of doxycycline, the DSB fragment generated by HO cleavage at the *MATa* locus was still processed to the faster mobility form; however, in the presence of doxycycline, the unprocessed 0.7-kb cut fragment persisted for more than 6 h and no *MATα* repaired products were detected (Fig. 4a). The cut fragments were also stabilized in the SSA assay in the absence of Sae2, even 24 h after I-SceI induction, confirming the block to resection (Fig. 4b). These results show the initial processing step is Sae2 dependent. Gene conversion occurs in the *sgs1Δ P_{tet}-SAE2* strain in the presence of doxycycline, which confirms Exo1 is sufficient for DSB processing in the absence of Sgs1 (Supplementary Fig. 5d).

Conclusions

We have identified the key factors that are essential for generating 3' single-stranded tails in DSB processing. In agreement with previous studies, the Mre11 complex and Sae2/CtIP have a critical role, but our findings suggest this function is limited to an early step in the reaction (Fig. 5). It is possible that the Mre11 complex and Sae2 fail to bind efficiently to the processed ends, preventing repeated cycles of cleavage. This initial cleavage event could be important to prevent binding by the Ku complex and subsequent repair by end joining, and

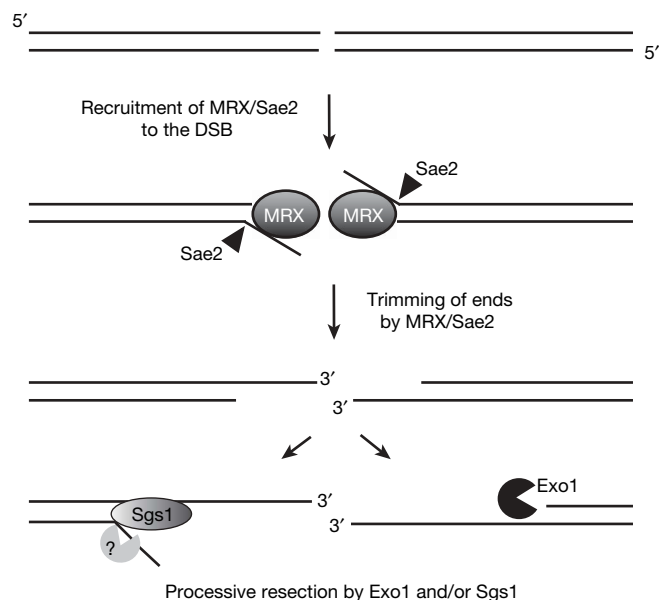


Figure 5 | Two-step mechanism for DSB resection. After a DSB is formed and recognized by the MRX complex, Sae2, in collaboration with MRX, trims the ends to create a minimally resected intermediate that is 'compatible' for processive resection by the 5'–3' exonucleolytic activity of Exo1 or Sgs1 helicase and a single-strand-specific nuclease. Cells lacking both the nuclease and the helicase activity accumulate the intermediates from MRX/Sae2 cleavage. Because *mre11Δ* and *sae2Δ* mutants still show DSB processing, Exo1 and Sgs1 must be able to access unprocessed DNA ends, but with reduced efficiency.

to provide a substrate for Sgs1, which shows higher affinity for 3' tailed substrates than blunt ends²⁶. The intermediate thus formed is rapidly processed by either the 5'–3' exonuclease activity of Exo1 or by Sgs1 helicase (Fig. 5). We assume Sgs1 functions with a single-strand-specific nuclease to remove the 5' strand. The essential nuclease, Dna2, is a candidate to function in resection because conditional mutations exhibit synthetic growth defects with *exo1Δ* and *sgs1Δ*²⁷. That Sgs1 acts at this early step is unexpected because genetic studies have implied a role for it after Rad51 action, and biochemical studies have shown the human BLM protein to dissociate D-loops and double Holliday junction intermediates^{28–30}. However, BLM localizes rapidly to sites of laser-light-induced DSBs, suggesting an early role in repair in addition to dissolution of recombination intermediates³¹. As all of the factors involved in resection in yeast are evolutionarily conserved, these findings suggest a general mechanism for DSB processing in eukaryotes.

METHODS SUMMARY

The yeast strains used were derivatives of W303. Details of their construction are provided in Methods. Physical analyses of mating-type switching and detection of single-strand intermediates at the *MAT* locus were as described previously^{5,21}. For the SSA assay, strains were grown to mid-log phase in medium containing lactate as a carbon source; I-SceI was induced for 1 h by addition of galactose to the growth medium. Samples were removed at the indicated times for DNA analysis. Genomic DNA was digested with NheI and EagI, and analysed by Southern blots using an *ADE2* probe. Full details are provided in Methods. For measurement of plating efficiency of strains containing the SSA reporter, cells were grown to exponential phase in YPLactate, then serially diluted (1:10) and spotted on solid rich medium with glucose (YPD) or galactose (YPRaf/Gal). The plates were incubated for 3 days and then scanned.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 May; accepted 1 August 2008.

Published online 21 September 2008.

1. Lee, S. E. et al. *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* **94**, 399–409 (1998).

2. Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).
3. Nelms, B. E., Maser, R. S., MacKay, J. F., Lagally, M. G. & Petrini, J. H. *In situ* visualization of DNA double-strand break repair in human fibroblasts. *Science* **280**, 590–592 (1998).
4. Krogh, B. O. & Symington, L. S. Recombination proteins in yeast. *Annu. Rev. Genet.* **38**, 233–271 (2004).
5. Llorente, B. & Symington, L. S. The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. *Mol. Cell. Biol.* **24**, 9682–9694 (2004).
6. Lobachev, K. S., Gordenin, D. A. & Resnick, M. A. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183–193 (2002).
7. Rattray, A. J., McGill, C. B., Shafer, B. K. & Strathern, J. N. Fidelity of mitotic double-strand-break repair in *Saccharomyces cerevisiae*: a role for *SAE2/COM1*. *Genetics* **158**, 109–122 (2001).
8. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The *Saccharomyces cerevisiae* Sae2 protein promotes resection and bridging of double strand break ends. *J. Biol. Chem.* **280**, 38631–38638 (2005).
9. Limbo, O. *et al.* Ctp1 is a cell-cycle-regulated protein that functions with Mre11 complex to control double-strand break repair by homologous recombination. *Mol. Cell* **28**, 134–146 (2007).
10. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
11. Tran, P. T., Erdeniz, N., Symington, L. S. & Liskay, R. M. EXO1-A multi-tasking eukaryotic nuclease. *DNA Repair (Amst.)* **3**, 1549–1559 (2004).
12. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The *Saccharomyces cerevisiae* Sae2 protein negatively regulates DNA damage checkpoint signalling. *EMBO Rep.* **7**, 212–218 (2006).
13. Hickson, I. D. RecQ helicases: caretakers of the genome. *Nature Rev. Cancer* **3**, 169–178 (2003).
14. Vaze, M. B. *et al.* Recovery from checkpoint-mediated arrest after repair of a double-strand break requires Srs2 helicase. *Mol. Cell* **10**, 373–385 (2002).
15. Mozlin, A. M., Fung, C. W. & Symington, L. S. Role of the *Saccharomyces cerevisiae* Rad51 paralogs in sister chromatid recombination. *Genetics* **178**, 113–126 (2008).
16. Ivanov, E. L., Sugawara, N., Fishman-Lobell, J. & Haber, J. E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **142**, 693–704 (1996).
17. Amundsen, S. K. & Smith, G. R. Interchangeable parts of the *Escherichia coli* recombination machinery. *Cell* **112**, 741–744 (2003).
18. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. & Rothstein, R. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol. Cell. Biol.* **14**, 8391–8398 (1994).
19. Watt, P. M., Louis, E. J., Borts, R. H. & Hickson, I. D. Sgs1: a eukaryotic homolog of *E. coli* RecQ that interacts with topoisomerase II *in vivo* and is required for faithful chromosome segregation. *Cell* **81**, 253–260 (1995).
20. Pan, X. *et al.* A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell* **124**, 1069–1081 (2006).
21. White, C. I. & Haber, J. E. Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *EMBO J.* **9**, 663–673 (1990).
22. Mullen, J. R., Kaliraman, V. & Brill, S. J. Bipartite structure of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. *Genetics* **154**, 1101–1114 (2000).
23. Jazayeri, A., Balestrini, A., Garner, E., Haber, J. E. & Costanzo, V. Mre11-Rad50-Nbs1-dependent processing of DNA breaks generates oligonucleotides that stimulate ATM activity. *EMBO J.* **27**, 1953–1962 (2008).
24. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell* **28**, 638–651 (2007).
25. Baron, U. & Bujard, H. Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol.* **327**, 401–421 (2000).
26. Bennett, R. J., Keck, J. L. & Wang, J. C. Binding specificity determines polarity of DNA unwinding by the Sgs1 protein of *S. cerevisiae*. *J. Mol. Biol.* **289**, 235–248 (1999).
27. Budd, M. E. *et al.* A network of multi-tasking proteins at the DNA replication fork preserves genome stability. *PLoS Genet* **1**, e61 (2005).
28. Gangloff, S., Soustelle, C. & Fabre, F. Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. *Nature Genet.* **25**, 192–194 (2000).
29. van Brabant, A. J. *et al.* Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* **39**, 14617–14625 (2000).
30. Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874 (2003).
31. Karmakar, P. *et al.* BLM is an early responder to DNA double-strand breaks. *Biochem. Biophys. Res. Commun.* **348**, 62–69 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Lichten, J. McCusker, A. Rattray and R. Rothstein for gifts of strains and plasmids, and W. K. Holloman and R. Rothstein for comments on the manuscript. This study was supported by a grant from the National Institutes of Health.

Author Contributions E.P.M. and L.S.S. designed the experiments and wrote the paper; E.P.M. performed the experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.S.S. (lss5@columbia.edu).

METHODS

Yeast strains and plasmids. The strains used were derived from W303 (ref. 32) and are listed in Supplementary Table 1. Strain LSY1092 was made by one-step gene replacement of W1588-4A with EcoRI/SphI-digested pMJ536 (a gift of M. Lichten) to replace the *SAE2* coding region with the *kanMX6* cassette³³. Strain LSY1975 was constructed by transformation of W1588-4C with a PCR fragment to generate a deletion allele of *SGS1*. PCR was performed on an *hphMX4* cassette plasmid, pAG32 (a gift of J. McCusker)³⁴. Details of the primers used for gene disruption and confirmation are available on request. To create strain LSY2041, a PCR fragment amplifying the tet-*SAE2* cassette from YAR1006 (gift of A. Rattray) was transformed into W1588-4C. A PCR fragment amplifying a 13-Myc tag from plasmid pFA6a-13Myc-*TRP1* (ref. 35) was transformed into LSY2041 to generate strain LSY2077. Strains containing the *ade2* direct repeat were generated by crossing LSY1430 (ref. 15) to strains within the laboratory collection to produce haploid progeny of the indicated genotypes.

Plasmids containing the wild-type or helicase-defective *SGS1* were provided by R. Rothstein. pEM-*SGS1* and pEM-*sgs1-K706A* were created by moving the *SGS1*-containing KpnI/SacI fragment from plasmids pSM100 and pRS415-*sgs1-K706A*, respectively, into the multiple cloning site of pRS316.

Media, growth conditions, and genetic methods. Media, growth conditions and genetic methods are as described previously³⁶. Hygromycin B (Sigma) to 300 $\mu\text{g ml}^{-1}$ was used for selection of the *hphMX4* cassette. Doxycycline (Sigma) was used to 20 $\mu\text{g ml}^{-1}$ final concentration in the conditional Sae2 experiments. G418 (Sigma) to 200 $\mu\text{g ml}^{-1}$ was used for selection of the *kanMX* cassettes.

Physical analysis of mating-type switching. Strains to be tested were transformed to Trp⁺ with HO-expressing plasmid, pFH800 (ref. 37). Trp⁺ transformants were grown in 5 ml of SC medium lacking tryptophan (SC-Trp) for 18 h. Cells were harvested, washed with water and used to inoculate 200 ml of SC-Raf-Trp. Cultures were grown to an optical density at 600 nm ($\text{OD}_{600\text{ nm}}$) of 0.3–0.4, a 50-ml sample was removed ($t = 0$ h) and galactose was added to the medium to a final concentration of 2%. One hour later, the cultures were harvested, washed and re-suspended in 200 ml of SC-Raf-Trp + 2% glucose. Samples (40 ml) were removed at 1-h intervals after induction for DNA analysis. Cells were harvested by centrifugation and the cell pellets were stored at -20°C . DNA was extracted, digested with StyI, and DNA fragments were separated by electrophoresis through 1% agarose gels. DNA fragments were transferred to nylon membranes and hybridized with a PCR fragment generated by amplification of *MAT* sequences distal to the HO-cut site (coordinates 201176–201580 on chromosome III sequence). For the conditional Sae2 experiments, the same protocol was followed except that before addition of galactose, the cells were pre-treated with 20 $\mu\text{g ml}^{-1}$ doxycycline for 6 h to ensure full repression of Sae2 expression.

To quantify the repair efficiency, we compared the intensity of the gene conversion product to the total intensity of the lane, using ImageJ (National Institutes of Health, USA). This ratio was then corrected for the HO cutting efficiency for each experiment. The cutting efficiency was quantified by using a set of primers that specifically annealed to Ya and Z2 sequences and amplified only the uncut substrate. PCR conditions for log-linear region amplification were used, and PCR was performed using as templates genomic DNA from $t = 0$ h and $t = 2$ h. In the same PCR reaction, a second set of primers with the same annealing temperature but unrelated target locus was used as loading control. The PCR products were analysed on agarose gels and the intensity of

the bands was quantified using ImageJ. For each time point, we calculated the ratio of uncut MAT/control, designated as x . The ratio of $x_{2\text{ h}}/x_{0\text{ h}}$ provided the percentage of uncut MATa that, when subtracted from 1, gave the cutting efficiency.

Physical analysis of single-strand annealing. Strains to be tested were grown in 5 ml of SC-Trp. Cells were harvested, washed with water and used to inoculate 200 ml of YPLactate (3%) medium. Cultures were grown to an $\text{OD}_{600\text{ nm}}$ of 0.3–0.4, a 50-ml sample was removed ($t = 0$ h) and 16.8 ml of 20% galactose were added to the medium. One hour later, the cultures were harvested, washed and resuspended in 200 ml of YPLactate + 2% glucose. Samples were removed at 1-h intervals after induction for DNA analysis as described above. DNA was digested with NheI/EagI, and DNA fragments were separated by electrophoresis through 0.8% agarose gels. DNA fragments were transferred to nylon membranes and hybridized with a 3.7-kb BglII fragment of the *ADE2* locus. The final SSA product was quantified by the ratio of the intensity of the SSA product to the whole intensity of the lane. For the conditional Sae2 experiments, the same protocol was followed, except that before addition of galactose, the cells were pre-treated with 20 $\mu\text{g ml}^{-1}$ doxycycline for 6 h.

Dot-blot analysis. DNA samples were applied to nylon membranes in the native or denatured state. Total genomic DNA was used for the native and denatured samples, 1 μg versus 0.1 μg , respectively, after adjusting the concentration to $10\times\text{SSC}$. Samples were transferred to nylon membranes using a dot-blot apparatus. The membrane was ultraviolet cross-linked, and the blots were hybridized with the *ADE2* probe. The signal of the dots was quantified by scanning the optic density of each dot using ImageJ.

Alkaline electrophoresis. ssDNA intermediates were analysed by alkaline gel electrophoresis as described²¹, and the blots hybridized with single-stranded probes complementary to the 5' or 3' strand. The probes were obtained by *in vitro* transcription using Epicentre Riboscribe T7 synthesis system and plasmids pEM-MAT or pEM-TAM as templates. The plasmids were obtained by cloning a *MAT* locus PCR fragment (coordinates 201176–201580 on chromosome III sequence) into pGEM-T Easy (Promega) in either orientation. To quantify the r2 band over time in different strain backgrounds, we plotted the ratio of intensity of the r2 band to the intensity of an unrelated locus as a loading control.

Western blot analysis. Whole-cell extracts were analysed by SDS–polyacrylamide gel electrophoresis, and Sae-Myc was detected using anti-c-myc antibody (Sigma).

32. Zou, H. & Rothstein, R. Holliday junctions accumulate in replication mutants via a RecA homolog-independent mechanism. *Cell* **90**, 87–96 (1997).
33. Borde, V., Wu, T. C. & Lichten, M. Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 4832–4842 (1999).
34. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
35. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
36. Sherman, F., Fink, G. & Hicks, J. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, 1986).
37. Nickoloff, J. A., Singer, J. D., Hoekstra, M. F. & Heffron, F. Double-strand breaks stimulate alternative mechanisms of recombination repair. *J. Mol. Biol.* **207**, 527–541 (1989).

ARTICLES

Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing

Eleni P. Mimitou¹ & Lorraine S. Symington¹

DNA ends exposed after introduction of double-strand breaks (DSBs) undergo 5'–3' nucleolytic degradation to generate single-stranded DNA, the substrate for binding by the Rad51 protein to initiate homologous recombination. This process is poorly understood in eukaryotes, but several factors have been implicated, including the Mre11 complex (Mre11–Rad50–Xrs2/NBS1), Sae2/CtIP/Ctp1 and Exo1. Here we demonstrate that yeast Exo1 nuclease and Sgs1 helicase function in alternative pathways for DSB processing. Novel, partially resected intermediates accumulate in a double mutant lacking Exo1 and Sgs1, which are poor substrates for homologous recombination. The early processing step that generates partly resected intermediates is dependent on Sae2. When Sae2 is absent, in addition to Exo1 and Sgs1, unprocessed DSBs accumulate and homology-dependent repair fails. These results suggest a two-step mechanism for DSB processing during homologous recombination. First, the Mre11 complex and Sae2 remove a small oligonucleotide(s) from the DNA ends to form an early intermediate. Second, Exo1 and/or Sgs1 rapidly process this intermediate to generate extensive tracts of single-stranded DNA that serve as substrate for Rad51.

DSBs are potentially lethal lesions that can occur spontaneously during normal cell metabolism or by treatment of cells with DNA-damaging agents. If unrepaired or repaired inappropriately, DSBs can lead to mutagenic events such as chromosome loss, deletions, duplications or translocations. The Mre11 complex (Mre11–Rad50–Xrs2/NBS1) is rapidly recruited to DSBs, signals checkpoint activation through Tel1/ATM kinase and regulates 5'–3' resection of the DNA ends^{1–3}. The intrinsic Mre11 nuclease activity is essential for processing ends blocked by covalent adducts, such as Spo11 or hairpin-capped ends, but appears to be less important for processing HO endonuclease-induced DSBs^{4–7}. Sae2/CtIP/Ctp1 interacts with the Mre11 complex, and Sae2 deficiency is phenotypically similar to loss of the Mre11 nuclease activity^{6–10}. The functions of the Mre11 complex and Sae2 in processing DSBs are partly redundant with the 5'–3' double-strand exonuclease, Exo1 (ref. 11). In the absence of Exo1 and Sae2, 5'–3' resection of DSBs is reduced, but resection and homologous recombination still occur, suggesting the existence of at least one other factor in DSB processing¹². Here we demonstrate a function for the *Saccharomyces cerevisiae* RecQ helicase homologue, Sgs1, in DSB processing. As the Sgs1 homologue in humans, BLM, maintains genome integrity and guards against cancer predisposition¹³, our findings suggest that some of the defects observed in individuals with Bloom's syndrome could be due to altered processing of DSBs.

Sae2 and Exo1 function at different steps in resection

We developed a single-strand annealing assay (SSA), which is inherently sensitive to defects in DSB resection, to identify genes required for this process¹⁴. The assay uses tandem repeats of the *ade2* gene, one of which contains a recognition site for the I-SceI endonuclease (Fig. 1a)¹⁵. The strain also harbours a *GAL1p-I-SCEI* fusion to provide regulated expression of the nuclease. After induction of I-SceI, homology-dependent repair can occur by gene conversion directed by the uncut allele, or by SSA⁴. SSA requires extensive 5'–3' degradation (about 7 kilobases (kb) in this case) to expose complementary single-stranded DNA (ssDNA) followed by Rad52-dependent

annealing of the repeats. After trimming residual ssDNA tails and gap filling, recombinants form with loss of one of the repeats and the intervening sequence. In the *rad51Δ* mutant, gene conversion is eliminated, leaving repair of the DSB by SSA as the alternative mode¹⁶. The efficiency of SSA can be assessed by the plating efficiency of strains on medium containing galactose (YP_{Raf}/Gal) to induce constitutive expression of I-SceI. Changes in the kinetics of SSA can be evaluated by Southern blot analysis of NheI/EagI-digested genomic DNA.

The plating efficiency of the *rad51Δ* strain on YP_{Raf}/Gal was the same as on medium containing glucose (YPD), and Ade⁺ recombinants were evident by formation of white colonies (Fig. 1b). By physical analysis, DSB fragments appeared 1 h after inducing I-SceI expression and were then processed to form the SSA product visible at 2 h and accumulating to 80% of the total DNA after 24 h (Fig. 1c). 5'–3' resection past the EagI and NheI sites results in loss of the cut fragments. Therefore, increased stability of the cut fragments indicates a delay or reduced rate of processing. Although the DSB fragments produced in the *rad51Δ exo1Δ* mutant were processed with similar kinetics as in the *rad51Δ* mutant, there was a delay in SSA product formation and reduced yield. In the *rad51Δ sae2Δ* strain, the DSB fragments persisted for longer, but the deletion products still appeared at 2 h and accumulated to 50% of the total. The phenotypes of *sae2Δ* and *exo1Δ* mutants, stability of cut fragments and delay in product formation, respectively, were additive in the *rad51Δ exo1Δ sae2Δ* mutant strain, resulting in a substantial defect in SSA and a twofold reduction in plating efficiency on YP_{Raf}/Gal plates (Fig. 1). To characterize the formation of ssDNA, DNA samples were applied to nylon membranes in the native or denatured state, and tested for ability to hybridize to a probe adjacent to the DSB. This analysis revealed similar amounts of ssDNA adjacent to the break site in *rad51Δ exo1Δ* and *rad51Δ* strains at early time points, whereas a decrease was observed in the *rad51Δ sae2Δ* mutant (Fig. 1d). ssDNA persisted in the *rad51Δ exo1Δ* mutant, consistent with the observed delay in SSA product formation. These results are in line with a model in which Sae2 starts DSB processing to produce an early

¹Department of Microbiology, Columbia University Medical Center, 701 West 168th Street, New York, New York 10032, USA.

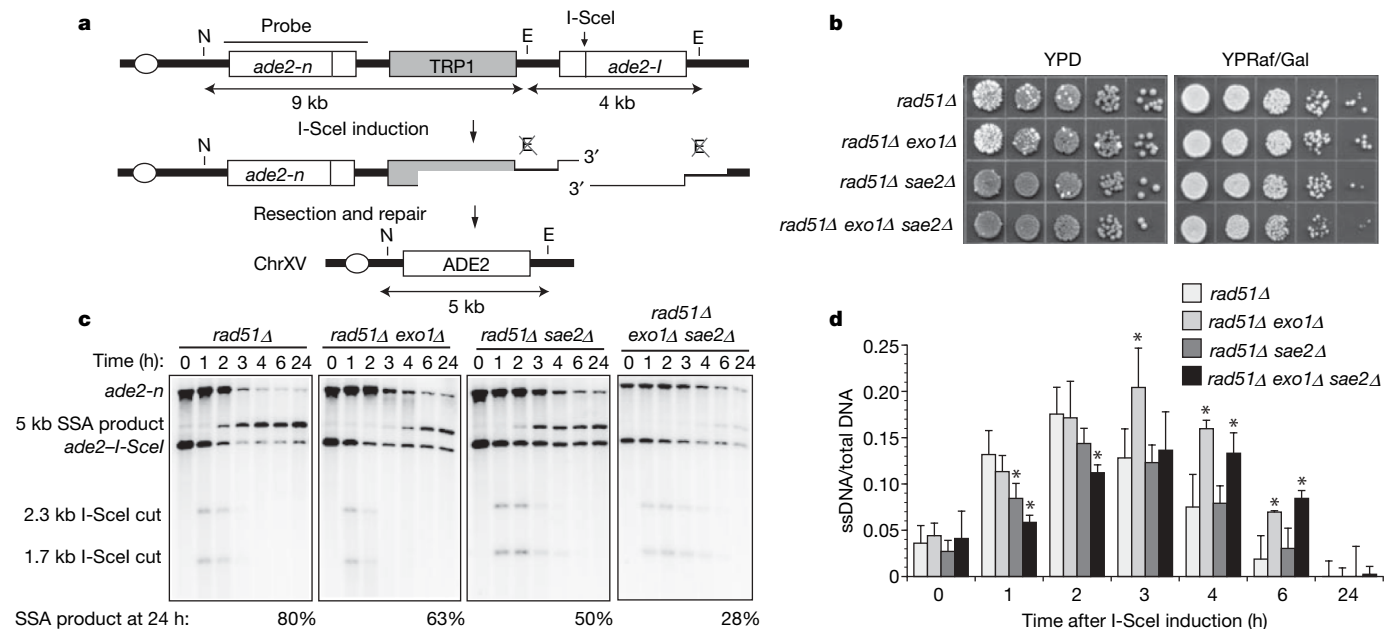


Figure 1 | Sae2 and Exo1 function at different steps in DSB resection.

a, Map of the chromosome XV region containing the *ade2* direct repeat. The *ade2-I* allele is cleaved by I-SceI to yield the 2.3- and 1.7-kb cut fragments that disappear over time, concomitant with repair to the 5-kb SSA product. **b**, SSA efficiency as indicated by sectoring on YPD plates and survival on YPRaf/Gal plates. Ade⁺ (white) colonies arise from *ade2* (red) colonies by SSA during growth on YPD owing to leaky expression of I-SceI. **c**, SSA

physical assay: monitoring of resection and repair of the I-SceI-induced DSB by Southern blot analysis of NheI/EagI-digested genomic DNA.

d, Quantification of the ADE2 hybridization signal on native versus denatured DNA samples bound on nylon membrane. The means from multiple experiments from three independent inductions are presented, the error bars indicate standard deviations; **P* < 0.02 (unpaired *t*-test).

intermediate that is rapidly matured by the processive nucleolytic activity of Exo1. In the absence of Sae2, a smaller subset of DSBs is acted on directly by Exo1 yielding SSA products at the same time, but with lower yield than found in the presence of Sae2.

DSB processing requires Sgs1 in the absence of Exo1

The presence of SSA products in the *exo1Δ sae2Δ* mutant suggests at least one other activity is operative in DSB processing. The missing

factor is not Mre11 because SSA products still form in *mre11Δ*, *mre11Δ exo1Δ*, and *mre11Δ exo1Δ sae2Δ* mutants (Supplementary Fig. 1). The nuclease-defective *mre11-H125N* mutant was slightly less defective than *sae2Δ*, by itself or with *exo1Δ* (data not shown). One possible mechanism for DSB processing would involve a helicase in conjunction with a single-strand-specific endo- or exonuclease. In *Escherichia coli*, the RecQ 3'–5' helicase and RecJ 5'–3' exonuclease function in DSB resection in the absence of the dominant RecBCD

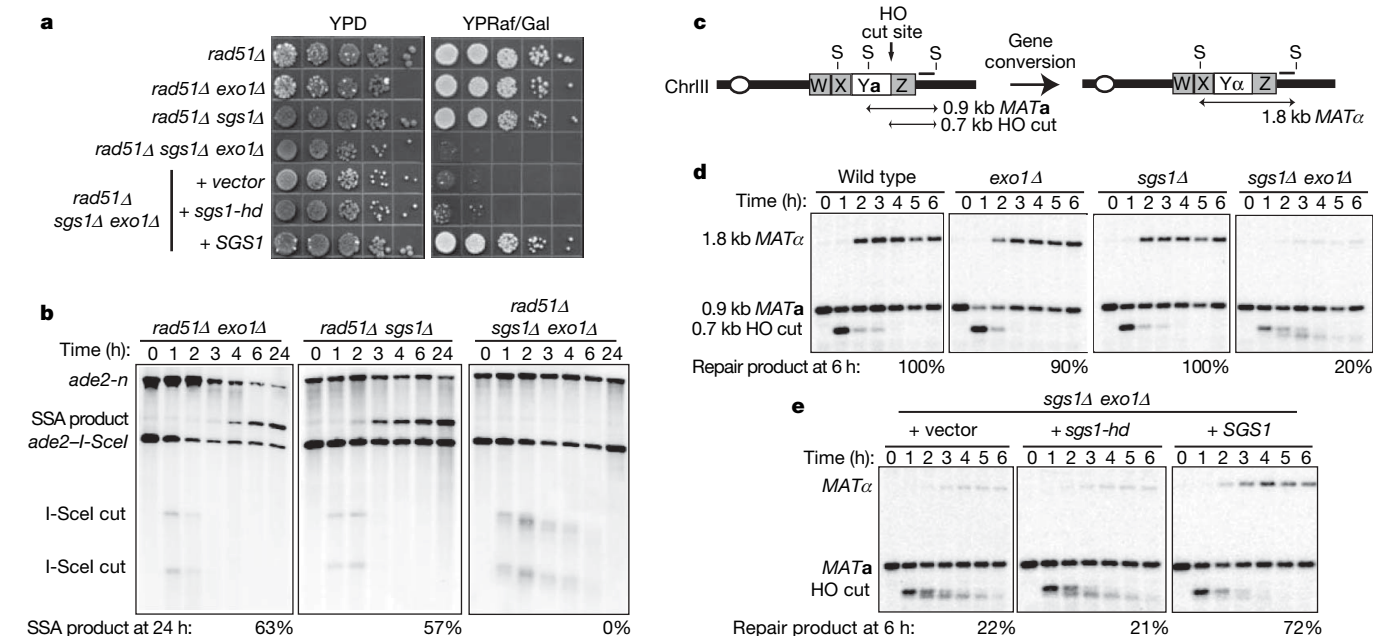


Figure 2 | Sgs1 and Exo1 function redundantly in DSB resection. **a**, SSA plating efficiency as described in Fig. 1b. **b**, SSA physical assay as described in Fig. 1c. **c**, Representation of MAT switching: the 0.9-kb MATa StyI fragment is cleaved by HO, resulting in a 0.7-kb cut fragment that disappears over time

concomitant with repair to a 1.8-kb MATα fragment. **d**, MAT switching assay: HO cut fragment processing and gene conversion efficiency as monitored by Southern blot analysis of StyI-digested genomic DNA. **e**, MAT switching assay on *sgs1Δ exo1Δ* cultures complemented with mutant or wild-type SGS1.

processing activity¹⁷. Therefore, we considered the possibility that the RecQ homologue of *S. cerevisiae*, Sgs1 (refs 18, 19), might fulfil a similar redundant function. The *rad51Δ sgs1Δ* mutant showed a delay in SSA (Fig. 2 and Supplementary Fig. 2); however, the *rad51Δ exo1Δ sgs1Δ* strain was completely deficient, as evidenced by the greater than 1,000-fold decrease in plating efficiency on YPRaf/Gal medium (Fig. 2a). Furthermore, there was no detectable deletion product by Southern blot analysis, even 24 h after I-SceI induction (Fig. 2b). Notably, in the *rad51Δ exo1Δ sgs1Δ* mutant novel intermediates were observed with slightly faster mobility than the DSB cut fragments. These became more smeared over time, suggesting very slow processing. We were unable to analyse the *sae2Δ sgs1Δ* double mutant because of a synthetic growth defect²⁰.

The SSA assay requires more extensive resection than is likely to be required for Rad51-dependent strand invasion. To determine whether the partly processed intermediates that accumulate in *exo1Δ sgs1Δ* strains are proficient for gene conversion, we analysed DSB processing in *RAD51* derivatives by a mating-type (*MAT*) switching assay. In this assay, DNA samples are digested with StyI to detect HO endonuclease cleavage and gene conversion products by Southern blot analysis (Fig. 2c)²¹. The *exo1Δ* and *sae2Δ* mutants exhibited slight defects in the efficiency of repair, whereas the *sgs1Δ* single mutant appeared normal (Fig. 2d and Supplementary Fig. 3). In the *exo1Δ sgs1Δ* double mutant, the DSB fragment was processed to yield an intermediate with faster mobility than the cut fragment. This persisted for several hours, as we had observed in the SSA assay. Despite the defect in DSB processing, the 1.8-kb gene-conversion product was still detected, although with reduced efficiency (20%) compared with wild type and the single mutants. This suggests very short ssDNA tails are sufficient to initiate Rad51-dependent gene conversion. To determine whether the helicase activity of Sgs1 is important for DSB processing, plasmids expressing wild-type or helicase-defective alleles of *SGS1* were used to transform the *exo1Δ sgs1Δ* double mutant, and the resulting strains tested for *MAT* switching or SSA proficiency²². As anticipated, the *SGS1* plasmid rescued the SSA and gene conversion defects of the *sgs1Δ exo1Δ* mutant, but the helicase-defective *sgs1-K706A* allele failed to complement *sgs1Δ* in these assays (Fig. 2). Previous studies have shown the steady-state level of the Sgs1-K706A protein to be similar to wild type²². Therefore, the helicase activity is required for DSB processing by Sgs1. The lack of full complementation by *SGS1* in the *MAT*-switching assay could be due to plasmid loss in some of the cells in the population.

Resection is attenuated in the *exo1Δ sgs1Δ* strain

To measure resection of 5' strands directly, we analysed HO-induced DSB fragments by restriction digestion and alkaline gel electrophoresis (Fig. 3a). Processing of the HO-induced break renders the DNA single-stranded and resistant to digestion with restriction enzymes, resulting in a ladder of bands of higher molecular mass²¹. By using RNA probes specific for the 3' and 5' strands to the right of the HO cut site, resection of the strands can be monitored. A *rad51Δ* background was used for this analysis to prevent engagement of the intermediates in repair.

The probe specific for 3' strands detected extensive resection of the HO-induced break in the *rad51Δ* strain (Fig. 3b). In the *rad51Δ exo1Δ* mutant, the first two ssDNA intermediates were formed with normal kinetics. However, the products indicative of extensive resection were delayed, resulting in accumulation of the r2 intermediate, suggesting Sgs1-dependent resection is less processive (Supplementary Fig. 4)⁵. The extent and timing of resection in the *rad51Δ sgs1Δ* mutant were similar to the *rad51Δ* strain. In contrast, only a faint band corresponding to resection beyond the first StyI site was detected in the *rad51Δ exo1Δ sgs1Δ* strain. Detection of a discrete band corresponding to degradation beyond the first StyI site and failure to detect products smaller than the cut fragment suggest the 3' ends are intact in the *rad51Δ exo1Δ sgs1Δ* mutant. This was confirmed by using a probe specific for the 5' strand. No intermediates

were detected in the *rad51Δ exo1Δ* or *rad51Δ sgs1Δ* mutants, presumably because of the rapid degradation of 5' ends. However, a smeared product was detected with faster migration than the 700-base-pair (bp) cut fragment in the *rad51Δ exo1Δ sgs1Δ* mutant, consistent with the severe resection defect. Interestingly, the smear does not occur immediately below the cut fragment, suggesting there is an initial endonuclease cleavage step that removes about 50–100 nucleotides from the 5' end.

Two-step mechanism for DSB processing

The results presented above suggest that after formation of a DSB, the ends are trimmed by an endonuclease to an intermediate form, which is subsequently processed by Exo1 or Sgs1. The Mre11 complex and Sae2 are likely to be involved in the initial processing step because DSB fragments are stabilized in *mre11Δ* and *sae2Δ* mutants, and recent studies have shown Mre11-dependent processing of DSBs to form short ssDNA oligonucleotides²³. Furthermore, Sae2 has an endonuclease activity that is stimulated by the Mre11 complex²⁴. We attempted to generate *exo1Δ sgs1Δ mre11Δ* or *exo1Δ sgs1Δ sae2Δ* triple mutants, but these were inviable (Supplementary Fig. 5a, b); the *exo1Δ sgs1Δ mre11-H125N* triple mutant was also inviable. A conditional *sae2* allele was created in the *exo1Δ sgs1Δ* background by expressing endogenous *SAE2* from the *P_{tet}* promoter,

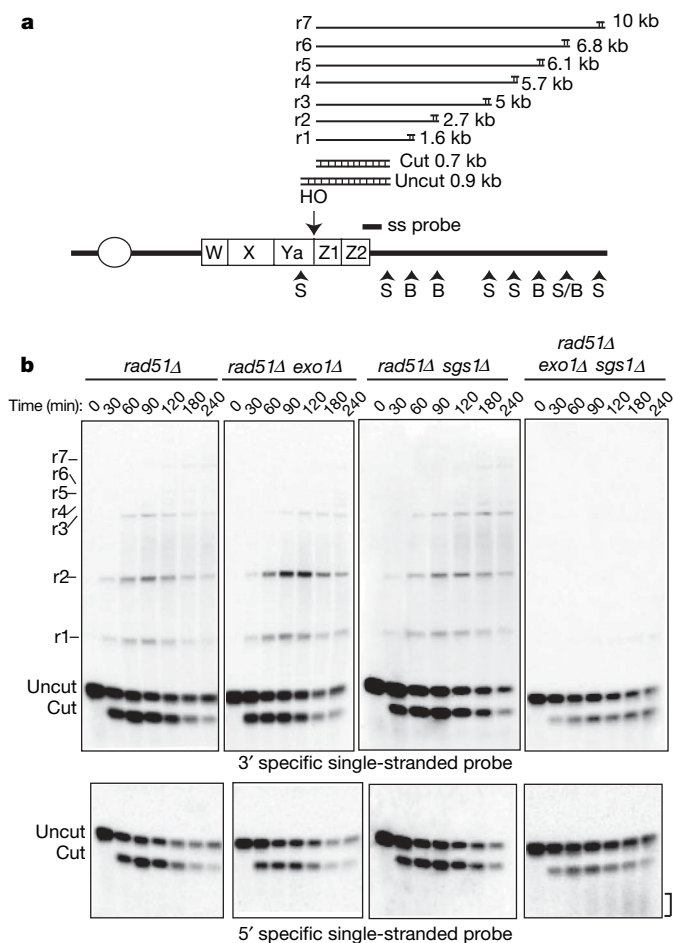


Figure 3 | Single-stranded intermediates fail to form in the absence of Exo1 and Sgs1. **a**, Representation of the method used to detect single-stranded intermediates after 5'–3' resection of the *MAT* locus, showing the positions of the HO cut site, the probe and the StyI (S) and BstXI (B) sites. Processing of the HO-induced break renders the DNA single-stranded and resistant to digestion with restriction enzymes, giving rise to distinct sets of fragments²¹. **b**, Alkaline electrophoresis of StyI/BstXI-digested genomic DNA: ssDNA intermediates formed after the resection of the HO-induced break were detected by using 3'- or 5'-specific riboprobes. The bracket indicates the smeared 5'-terminated strands observed in the *rad51Δ exo1Δ sgs1Δ* mutant.

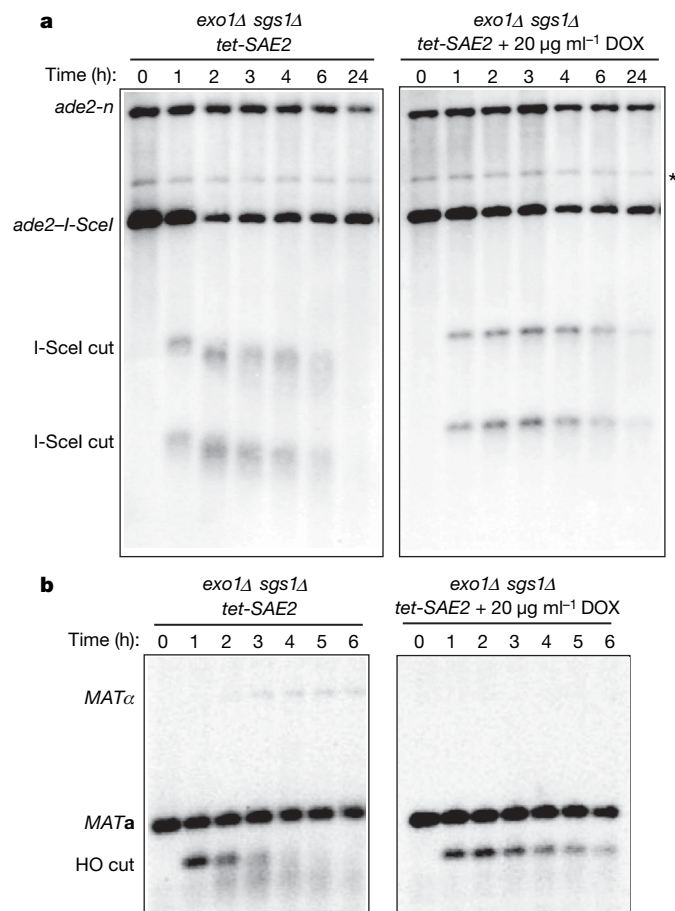


Figure 4 | Sae2 is required for creating the minimally resected intermediates. SSA assay (a) and mating-type switching assay (b) in cells treated or untreated with doxycycline before I-SceI or HO expression. The asterisk refers to a cross-hybridizing band, which is present at time 0 h; if it were due to recombinants in the population, it would allow survival of cells on YPRaf/Gal medium.

this strain also expresses the tetracycline-controlled trans-activator (tTA) to repress transcription of *SAE2* by growth in the presence of doxycycline²⁵. The *exo1Δ sgs1Δ P_{tet}-SAE2* strain was viable in the absence of doxycycline, indicating normal function. By western blot analysis, the Sae2 protein was depleted 4 h after addition of doxycycline to the growth medium (Supplementary Fig. 5c). In the absence of doxycycline, the DSB fragment generated by HO cleavage at the *MATa* locus was still processed to the faster mobility form; however, in the presence of doxycycline, the unprocessed 0.7-kb cut fragment persisted for more than 6 h and no *MATα* repaired products were detected (Fig. 4a). The cut fragments were also stabilized in the SSA assay in the absence of Sae2, even 24 h after I-SceI induction, confirming the block to resection (Fig. 4b). These results show the initial processing step is Sae2 dependent. Gene conversion occurs in the *sgs1Δ P_{tet}-SAE2* strain in the presence of doxycycline, which confirms Exo1 is sufficient for DSB processing in the absence of Sgs1 (Supplementary Fig. 5d).

Conclusions

We have identified the key factors that are essential for generating 3' single-stranded tails in DSB processing. In agreement with previous studies, the Mre11 complex and Sae2/CtIP have a critical role, but our findings suggest this function is limited to an early step in the reaction (Fig. 5). It is possible that the Mre11 complex and Sae2 fail to bind efficiently to the processed ends, preventing repeated cycles of cleavage. This initial cleavage event could be important to prevent binding by the Ku complex and subsequent repair by end joining, and

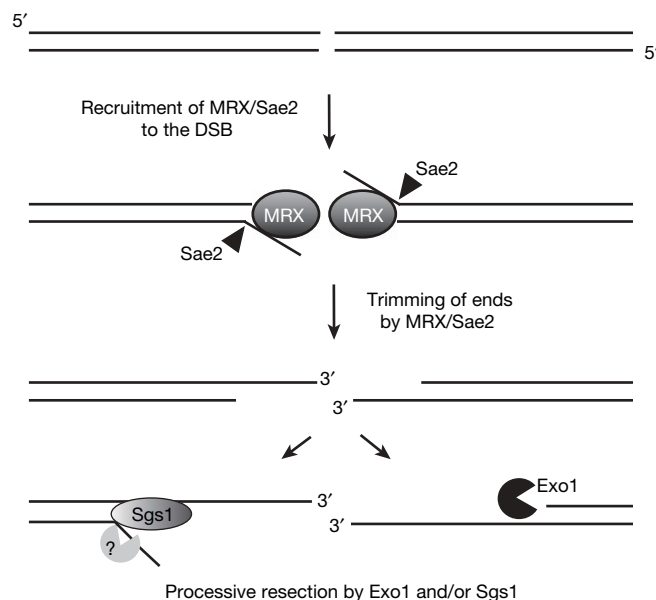


Figure 5 | Two-step mechanism for DSB resection. After a DSB is formed and recognized by the MRX complex, Sae2, in collaboration with MRX, trims the ends to create a minimally resected intermediate that is 'compatible' for processive resection by the 5'–3' exonucleolytic activity of Exo1 or Sgs1 helicase and a single-strand-specific nuclease. Cells lacking both the nuclease and the helicase activity accumulate the intermediates from MRX/Sae2 cleavage. Because *mre11Δ* and *sae2Δ* mutants still show DSB processing, Exo1 and Sgs1 must be able to access unprocessed DNA ends, but with reduced efficiency.

to provide a substrate for Sgs1, which shows higher affinity for 3' tailed substrates than blunt ends²⁶. The intermediate thus formed is rapidly processed by either the 5'–3' exonuclease activity of Exo1 or by Sgs1 helicase (Fig. 5). We assume Sgs1 functions with a single-strand-specific nuclease to remove the 5' strand. The essential nuclease, Dna2, is a candidate to function in resection because conditional mutations exhibit synthetic growth defects with *exo1Δ* and *sgs1Δ*²⁷. That Sgs1 acts at this early step is unexpected because genetic studies have implied a role for it after Rad51 action, and biochemical studies have shown the human BLM protein to dissociate D-loops and double Holliday junction intermediates^{28–30}. However, BLM localizes rapidly to sites of laser-light-induced DSBs, suggesting an early role in repair in addition to dissolution of recombination intermediates³¹. As all of the factors involved in resection in yeast are evolutionarily conserved, these findings suggest a general mechanism for DSB processing in eukaryotes.

METHODS SUMMARY

The yeast strains used were derivatives of W303. Details of their construction are provided in Methods. Physical analyses of mating-type switching and detection of single-strand intermediates at the *MAT* locus were as described previously^{5,21}. For the SSA assay, strains were grown to mid-log phase in medium containing lactate as a carbon source; I-SceI was induced for 1 h by addition of galactose to the growth medium. Samples were removed at the indicated times for DNA analysis. Genomic DNA was digested with NheI and EagI, and analysed by Southern blots using an *ADE2* probe. Full details are provided in Methods. For measurement of plating efficiency of strains containing the SSA reporter, cells were grown to exponential phase in YPLactate, then serially diluted (1:10) and spotted on solid rich medium with glucose (YPD) or galactose (YPRaf/Gal). The plates were incubated for 3 days and then scanned.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 13 May; accepted 1 August 2008.

Published online 21 September 2008.

1. Lee, S. E. et al. *Saccharomyces* Ku70, mre11/rad50 and RPA proteins regulate adaptation to G2/M arrest after DNA damage. *Cell* **94**, 399–409 (1998).

2. Lisby, M., Barlow, J. H., Burgess, R. C. & Rothstein, R. Choreography of the DNA damage response: spatiotemporal relationships among checkpoint and repair proteins. *Cell* **118**, 699–713 (2004).
3. Nelms, B. E., Maser, R. S., MacKay, J. F., Lagally, M. G. & Petrini, J. H. *In situ* visualization of DNA double-strand break repair in human fibroblasts. *Science* **280**, 590–592 (1998).
4. Krogh, B. O. & Symington, L. S. Recombination proteins in yeast. *Annu. Rev. Genet.* **38**, 233–271 (2004).
5. Llorente, B. & Symington, L. S. The Mre11 nuclease is not required for 5' to 3' resection at multiple HO-induced double-strand breaks. *Mol. Cell. Biol.* **24**, 9682–9694 (2004).
6. Lobachev, K. S., Gordenin, D. A. & Resnick, M. A. The Mre11 complex is required for repair of hairpin-capped double-strand breaks and prevention of chromosome rearrangements. *Cell* **108**, 183–193 (2002).
7. Rattray, A. J., McGill, C. B., Shafer, B. K. & Strathern, J. N. Fidelity of mitotic double-strand-break repair in *Saccharomyces cerevisiae*: a role for *SAE2/COM1*. *Genetics* **158**, 109–122 (2001).
8. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The *Saccharomyces cerevisiae* Sae2 protein promotes resection and bridging of double strand break ends. *J. Biol. Chem.* **280**, 38631–38638 (2005).
9. Limbo, O. *et al.* Ctp1 is a cell-cycle-regulated protein that functions with Mre11 complex to control double-strand break repair by homologous recombination. *Mol. Cell* **28**, 134–146 (2007).
10. Sartori, A. A. *et al.* Human CtIP promotes DNA end resection. *Nature* **450**, 509–514 (2007).
11. Tran, P. T., Erdeniz, N., Symington, L. S. & Liskay, R. M. EXO1-A multi-tasking eukaryotic nuclease. *DNA Repair (Amst.)* **3**, 1549–1559 (2004).
12. Clerici, M., Mantiero, D., Lucchini, G. & Longhese, M. P. The *Saccharomyces cerevisiae* Sae2 protein negatively regulates DNA damage checkpoint signalling. *EMBO Rep.* **7**, 212–218 (2006).
13. Hickson, I. D. RecQ helicases: caretakers of the genome. *Nature Rev. Cancer* **3**, 169–178 (2003).
14. Vaze, M. B. *et al.* Recovery from checkpoint-mediated arrest after repair of a double-strand break requires Srs2 helicase. *Mol. Cell* **10**, 373–385 (2002).
15. Mozlin, A. M., Fung, C. W. & Symington, L. S. Role of the *Saccharomyces cerevisiae* Rad51 paralogs in sister chromatid recombination. *Genetics* **178**, 113–126 (2008).
16. Ivanov, E. L., Sugawara, N., Fishman-Lobell, J. & Haber, J. E. Genetic requirements for the single-strand annealing pathway of double-strand break repair in *Saccharomyces cerevisiae*. *Genetics* **142**, 693–704 (1996).
17. Amundsen, S. K. & Smith, G. R. Interchangeable parts of the *Escherichia coli* recombination machinery. *Cell* **112**, 741–744 (2003).
18. Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. & Rothstein, R. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol. Cell. Biol.* **14**, 8391–8398 (1994).
19. Watt, P. M., Louis, E. J., Borts, R. H. & Hickson, I. D. Sgs1: a eukaryotic homolog of *E. coli* RecQ that interacts with topoisomerase II *in vivo* and is required for faithful chromosome segregation. *Cell* **81**, 253–260 (1995).
20. Pan, X. *et al.* A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell* **124**, 1069–1081 (2006).
21. White, C. I. & Haber, J. E. Intermediates of recombination during mating type switching in *Saccharomyces cerevisiae*. *EMBO J.* **9**, 663–673 (1990).
22. Mullen, J. R., Kaliraman, V. & Brill, S. J. Bipartite structure of the Sgs1 DNA helicase in *Saccharomyces cerevisiae*. *Genetics* **154**, 1101–1114 (2000).
23. Jazayeri, A., Balestrini, A., Garner, E., Haber, J. E. & Costanzo, V. Mre11-Rad50-Nbs1-dependent processing of DNA breaks generates oligonucleotides that stimulate ATM activity. *EMBO J.* **27**, 1953–1962 (2008).
24. Lengsfeld, B. M., Rattray, A. J., Bhaskara, V., Ghirlando, R. & Paull, T. T. Sae2 is an endonuclease that processes hairpin DNA cooperatively with the Mre11/Rad50/Xrs2 complex. *Mol. Cell* **28**, 638–651 (2007).
25. Baron, U. & Bujard, H. Tet repressor-based system for regulated gene expression in eukaryotic cells: principles and advances. *Methods Enzymol.* **327**, 401–421 (2000).
26. Bennett, R. J., Keck, J. L. & Wang, J. C. Binding specificity determines polarity of DNA unwinding by the Sgs1 protein of *S. cerevisiae*. *J. Mol. Biol.* **289**, 235–248 (1999).
27. Budd, M. E. *et al.* A network of multi-tasking proteins at the DNA replication fork preserves genome stability. *PLoS Genet* **1**, e61 (2005).
28. Gangloff, S., Soustelle, C. & Fabre, F. Homologous recombination is responsible for cell death in the absence of the Sgs1 and Srs2 helicases. *Nature Genet.* **25**, 192–194 (2000).
29. van Brabant, A. J. *et al.* Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* **39**, 14617–14625 (2000).
30. Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874 (2003).
31. Karmakar, P. *et al.* BLM is an early responder to DNA double-strand breaks. *Biochem. Biophys. Res. Commun.* **348**, 62–69 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Lichten, J. McCusker, A. Rattray and R. Rothstein for gifts of strains and plasmids, and W. K. Holloman and R. Rothstein for comments on the manuscript. This study was supported by a grant from the National Institutes of Health.

Author Contributions E.P.M. and L.S.S. designed the experiments and wrote the paper; E.P.M. performed the experiments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.S.S. (lss5@columbia.edu).

METHODS

Yeast strains and plasmids. The strains used were derived from W303 (ref. 32) and are listed in Supplementary Table 1. Strain LSY1092 was made by one-step gene replacement of W1588-4A with EcoRI/SphI-digested pMJ536 (a gift of M. Lichten) to replace the *SAE2* coding region with the *kanMX6* cassette³³. Strain LSY1975 was constructed by transformation of W1588-4C with a PCR fragment to generate a deletion allele of *SGS1*. PCR was performed on an *hphMX4* cassette plasmid, pAG32 (a gift of J. McCusker)³⁴. Details of the primers used for gene disruption and confirmation are available on request. To create strain LSY2041, a PCR fragment amplifying the tet-*SAE2* cassette from YAR1006 (gift of A. Rattray) was transformed into W1588-4C. A PCR fragment amplifying a 13-Myc tag from plasmid pFA6a-13Myc-*TRP1* (ref. 35) was transformed into LSY2041 to generate strain LSY2077. Strains containing the *ade2* direct repeat were generated by crossing LSY1430 (ref. 15) to strains within the laboratory collection to produce haploid progeny of the indicated genotypes.

Plasmids containing the wild-type or helicase-defective *SGS1* were provided by R. Rothstein. pEM-*SGS1* and pEM-*sgs1-K706A* were created by moving the *SGS1*-containing KpnI/SacI fragment from plasmids pSM100 and pRS415-*sgs1-K706A*, respectively, into the multiple cloning site of pRS316.

Media, growth conditions, and genetic methods. Media, growth conditions and genetic methods are as described previously³⁶. Hygromycin B (Sigma) to 300 $\mu\text{g ml}^{-1}$ was used for selection of the *hphMX4* cassette. Doxycycline (Sigma) was used to 20 $\mu\text{g ml}^{-1}$ final concentration in the conditional Sae2 experiments. G418 (Sigma) to 200 $\mu\text{g ml}^{-1}$ was used for selection of the *kanMX* cassettes.

Physical analysis of mating-type switching. Strains to be tested were transformed to Trp⁺ with HO-expressing plasmid, pFH800 (ref. 37). Trp⁺ transformants were grown in 5 ml of SC medium lacking tryptophan (SC-Trp) for 18 h. Cells were harvested, washed with water and used to inoculate 200 ml of SC-Raf-Trp. Cultures were grown to an optical density at 600 nm ($\text{OD}_{600\text{ nm}}$) of 0.3–0.4, a 50-ml sample was removed ($t = 0$ h) and galactose was added to the medium to a final concentration of 2%. One hour later, the cultures were harvested, washed and re-suspended in 200 ml of SC-Raf-Trp + 2% glucose. Samples (40 ml) were removed at 1-h intervals after induction for DNA analysis. Cells were harvested by centrifugation and the cell pellets were stored at -20°C . DNA was extracted, digested with StyI, and DNA fragments were separated by electrophoresis through 1% agarose gels. DNA fragments were transferred to nylon membranes and hybridized with a PCR fragment generated by amplification of *MAT* sequences distal to the HO-cut site (coordinates 201176–201580 on chromosome III sequence). For the conditional Sae2 experiments, the same protocol was followed except that before addition of galactose, the cells were pre-treated with 20 $\mu\text{g ml}^{-1}$ doxycycline for 6 h to ensure full repression of Sae2 expression.

To quantify the repair efficiency, we compared the intensity of the gene conversion product to the total intensity of the lane, using ImageJ (National Institutes of Health, USA). This ratio was then corrected for the HO cutting efficiency for each experiment. The cutting efficiency was quantified by using a set of primers that specifically annealed to Ya and Z2 sequences and amplified only the uncut substrate. PCR conditions for log-linear region amplification were used, and PCR was performed using as templates genomic DNA from $t = 0$ h and $t = 2$ h. In the same PCR reaction, a second set of primers with the same annealing temperature but unrelated target locus was used as loading control. The PCR products were analysed on agarose gels and the intensity of

the bands was quantified using ImageJ. For each time point, we calculated the ratio of uncut MAT/control, designated as x . The ratio of $x_{2\text{ h}}/x_{0\text{ h}}$ provided the percentage of uncut MATa that, when subtracted from 1, gave the cutting efficiency.

Physical analysis of single-strand annealing. Strains to be tested were grown in 5 ml of SC-Trp. Cells were harvested, washed with water and used to inoculate 200 ml of YPLactate (3%) medium. Cultures were grown to an $\text{OD}_{600\text{ nm}}$ of 0.3–0.4, a 50-ml sample was removed ($t = 0$ h) and 16.8 ml of 20% galactose were added to the medium. One hour later, the cultures were harvested, washed and resuspended in 200 ml of YPLactate + 2% glucose. Samples were removed at 1-h intervals after induction for DNA analysis as described above. DNA was digested with NheI/EagI, and DNA fragments were separated by electrophoresis through 0.8% agarose gels. DNA fragments were transferred to nylon membranes and hybridized with a 3.7-kb BglII fragment of the *ADE2* locus. The final SSA product was quantified by the ratio of the intensity of the SSA product to the whole intensity of the lane. For the conditional Sae2 experiments, the same protocol was followed, except that before addition of galactose, the cells were pre-treated with 20 $\mu\text{g ml}^{-1}$ doxycycline for 6 h.

Dot-blot analysis. DNA samples were applied to nylon membranes in the native or denatured state. Total genomic DNA was used for the native and denatured samples, 1 μg versus 0.1 μg , respectively, after adjusting the concentration to $10\times\text{SSC}$. Samples were transferred to nylon membranes using a dot-blot apparatus. The membrane was ultraviolet cross-linked, and the blots were hybridized with the *ADE2* probe. The signal of the dots was quantified by scanning the optic density of each dot using ImageJ.

Alkaline electrophoresis. ssDNA intermediates were analysed by alkaline gel electrophoresis as described²¹, and the blots hybridized with single-stranded probes complementary to the 5' or 3' strand. The probes were obtained by *in vitro* transcription using Epicentre Riboscribe T7 synthesis system and plasmids pEM-MAT or pEM-TAM as templates. The plasmids were obtained by cloning a *MAT* locus PCR fragment (coordinates 201176–201580 on chromosome III sequence) into pGEM-T Easy (Promega) in either orientation. To quantify the r2 band over time in different strain backgrounds, we plotted the ratio of intensity of the r2 band to the intensity of an unrelated locus as a loading control.

Western blot analysis. Whole-cell extracts were analysed by SDS–polyacrylamide gel electrophoresis, and Sae-Myc was detected using anti-c-myc antibody (Sigma).

32. Zou, H. & Rothstein, R. Holliday junctions accumulate in replication mutants via a RecA homolog-independent mechanism. *Cell* **90**, 87–96 (1997).
33. Borde, V., Wu, T. C. & Lichten, M. Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **19**, 4832–4842 (1999).
34. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
35. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
36. Sherman, F., Fink, G. & Hicks, J. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory, 1986).
37. Nickoloff, J. A., Singer, J. D., Hoekstra, M. F. & Heffron, F. Double-strand breaks stimulate alternative mechanisms of recombination repair. *J. Mol. Biol.* **207**, 527–541 (1989).

The formation and assembly of a typical star-forming galaxy at redshift $z \approx 3$

Daniel P. Stark¹, A. Mark Swinbank², Richard S. Ellis¹, Simon Dye³, Ian R. Smail² & Johan Richard¹

Recent studies of galaxies $\sim 2\text{--}3$ Gyr after the Big Bang have revealed large, rotating disks, similar to those of galaxies today^{1,2}. The existence of well-ordered rotation in galaxies during this peak epoch of cosmic star formation indicates that gas accretion is likely to be the dominant mode by which galaxies grow, because major mergers of galaxies would completely disrupt the observed velocity fields. But poor spatial resolution and sensitivity have hampered this interpretation; such studies have been limited to the largest and most luminous galaxies, which may have fundamentally different modes of assembly from those of more typical galaxies (which are thought to grow into the spheroidal components at the centres of galaxies similar to the Milky Way). Here we report observations of a typical star-forming galaxy at $z = 3.07$, with a linear resolution of ~ 100 parsecs. We find a well-ordered compact source in which molecular gas is being converted efficiently into stars, likely to be assembling a spheroidal bulge similar to those seen in spiral galaxies at the present day. The presence of undisrupted rotation may indicate that galaxies such as the Milky Way gain much of their mass by accretion rather than major mergers.

Resolved imaging spectroscopy provides a powerful route to probing the internal properties of young galaxies. Using adaptive optics, which corrects for the effects of atmospheric motion, internal velocity maps have been acquired for a small number of galaxies at redshifts of $z = 2\text{--}3$ ($2\text{--}3$ Gyr after the Big Bang) on scales of 0.15 arcsec (~ 1.3 kpc)^{1–3}. These early results do not yet reveal a clear picture. The largest galaxies studied show evidence for systematic rotation, consistent with forming spiral disks^{1,2}. However, smaller, more typical, galaxies observed at the same epoch show chaotic velocity fields consistent with more primitive systems³. The interpretation is still limited by poor spatial resolution: typical galaxies at such redshifts have angular sizes of order 0.25 arcsec and so, even with adaptive optics, the velocity field is poorly sampled. Moreover, the emission lines used to trace the velocity field are not uniformly bright across the galaxy, and co-adding pixels to improve the signal comes at the expense of a degradation in resolution.

An effective way to resolve these issues is to use gravitational lensing, which magnifies the images of distant galaxies^{4–7} and allows a much more detailed view than would otherwise be possible. To demonstrate the advances possible, we have combined laser guide star adaptive optics with gravitational lensing to study a star-forming galaxy (MACS J2135–0102 (ref. 8)) at $z = 3.075$ (seen 2.1 Gyr after the Big Bang). Detailed modelling of this system reveals that this galaxy is magnified in area 28 ± 3 times (and ~ 8 times along its major axis) by the combination of a foreground galaxy and galaxy cluster⁹ and that its properties are typical of Lyman-break galaxies at this redshift.

To map the redshifted [O III] $4,959$, $5,007$ and $H\beta$ $4,861$ Å emission lines, we employed the OSIRIS integral field spectrograph¹⁰ on the 10 m

Keck II telescope. Figure 1 shows a colour image of the target constructed from the optical Hubble Space Telescope image and the OSIRIS infrared spectroscopic data; clearly the two data sets are well matched in resolution. Using adaptive optics, together with the gravitational magnification, the image plane full-width at half-maximum corresponds to a physical scale of just ~ 120 pc in the source plane. Not only does this yield a roughly fivefold improvement in linear resolution in sampling the velocity field in comparison with earlier work, but the signal can be co-added in fainter regions while still maintaining an improved resolution over that achievable for unlensed sources.

To interpret our observations, we take into account the magnification caused by the foreground lens and reconstruct the resolved galaxy in the source plane, making use of the detailed mass model⁹, which defines the mapping from each location in the image plane. Figure 2 shows the resulting intensity and velocity maps in which binning has typically been undertaken with source-plane pixels of 0.02 arcsec, corresponding to a sampling of 150 pc. The rest-frame ultraviolet continuum distribution reveals an elongated morphology

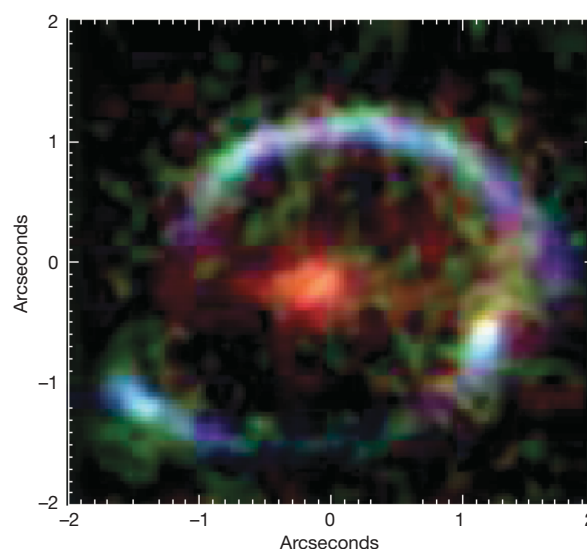


Figure 1 | Colour image of J2135–0102, a $z = 3.07$ galaxy magnified 28 ± 3 times into a near-complete Einstein ring⁹. This image combines a Hubble Space Telescope V_{606} image (blue) with near-infrared images from the Keck Laser Guide Star Adaptive Optics system in [O III] $5,007$ Å (green) and broadband K (red). The foreground lensing galaxy at $z = 0.7$ is the resolved red source at the centre of the image. The large extent of the ring compared with the image-plane resolution of our observations (0.13 arcsec) enables us to map the variation in spectral and dynamical properties of the background galaxy in fine detail.

¹Department of Astronomy, California Institute of Technology, Pasadena, California 91125, USA. ²Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK. ³School of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK.

with two components. Component 1 is the most highly magnified (a factor of 28 ± 3 times) and gives rise to the ring-like structure in the image plane. Component 2 extends just outside the caustic, giving rise to a lower magnification (a factor of 8.0 ± 0.9 times). These large magnifications enable us to study the internal properties of J2135–0102 in unprecedented detail. Fitting both the continuum and [O III] emission line distributions, we derive a half-light radius of 750 pc, fivefold smaller than the large galaxies studied by many earlier workers^{1,2} and much closer to the typical size of $z \approx 3$ Lyman break galaxies (LBGs)¹¹.

The most significant result revealed in Fig. 2 is the well-sampled, regular, bi-symmetric velocity field revealed by the [O III] emission line measures, which, together with the linewidths peaking at the centre of the galaxy, suggest a rotating system. But could the observed velocity pattern arise from a merging pair whose projected velocities are smoothed into an apparent rotation curve as a result of inadequate spatial resolution? This ambiguity, arising from the poorly sampled data typically available at high redshift, has plagued the interpretation of the first resolved data from distant galaxies^{1–4}, thereby precluding a robust test of the hypothesis that a large fraction of distant star-forming galaxies are undergoing major mergers¹². With our considerably improved resolution, this ambiguity is greatly reduced. The well-sampled data for J2135–0102 reveal a velocity gradient spanning many independent resolution elements even internally with component 1. Moreover, the projected velocity of component 2 is consistent with an extrapolation of the gradient across component 1, which would be highly coincidental if the components were separate systems. We therefore conclude that J2135–0102 it is much more likely to be a single system undergoing rotation.

Examining the velocity field in more detail, we extract a one-dimensional rotation curve along the major axis of the velocity field (at a position angle of 10° clockwise from the Δy axis). Using an arctan function to model the shape of the rotation curve, the observations suggest a circular velocity of $v_c \sin i = 55 \pm 7 \text{ km s}^{-1}$, where i is the inclination angle. We estimate $i = 55^\circ \pm 10^\circ$ from the surface photometry, implying a circular velocity of 67 km s^{-1} . As described above, component 2 shares the rotation, affecting only a minor $7 \pm 3 \text{ km s}^{-1}$ adjustment coincident with a 25 km s^{-1} increase in the linewidth. Using the rotation curve we derive a dynamical mass of $\sim 2 \times 10^9$ solar masses (M_\odot) within a radius of 1.8 kpc. Thus, in comparison with the stellar mass of $(6 \pm 2) \times 10^9 M_\odot$ (ref. 13) (assuming a Salpeter initial mass function), the central region seems to be baryon-dominated. Although J2135–0102 certainly represents the most convincing example of a rotating galaxy, significant random motions are present. The central [O III] velocity dispersion $\sigma_0 = 54 \pm 4 \text{ km s}^{-1}$ is comparable to the rotational velocity, and the ratio $v/\sigma_0 = 1.2 \pm 0.1$ is consistent with previous, more poorly sampled, observations of similar-sized galaxies at this redshift³. Any disk present in J2135–0102 is therefore most probably at an early stage of assembly (see Supplementary Information for further discussion).

Our OSIRIS data also yield the two-dimensional distribution of the nebular line H β , a valuable tracer of recent ($< 20 \text{ Myr}$) star formation. The integrated star formation rate derived from the H β flux is $40 \pm 5 M_\odot \text{ yr}^{-1}$ (uncorrected for extinction)¹⁴, which is consistent with that derived from the $24 \mu\text{m}$ flux¹³. The star formation rate density of $4.4 \pm 0.5 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$ is typical of that derived for a larger sample of ultraviolet-selected galaxies at $z \approx 2$ –3 (refs 15, 16) and comparable to that in local starbursts, in which activity is usually in compact, circumnuclear disks. The high rates of star formation observed in these galaxies result in significant stellar winds and supernovae ejecta, which are thought to contribute to the vigorous galactic scale outflows that have been observed in LBGs¹⁷. The blue-shifted ultraviolet interstellar absorption lines ($\Delta v = -400 \pm 100 \text{ km s}^{-1}$)⁸ are consistent with this interpretation.

We also show that the H β velocity field confirms the overall rotation. However, both the H β and CO emission¹³ are not co-spatial with

the centre of the galaxy but seem to peak within component 2. Moreover, the CO velocity and linewidth are in good agreement with those seen in [O III] for this component, indicating that almost all of the cold gas (which acts a reservoir for future star formation) may reside in a region less than 1 kpc across. We use additional spatially resolved ESO VLT SINFONI observations of [O II], [O III] and H β (see

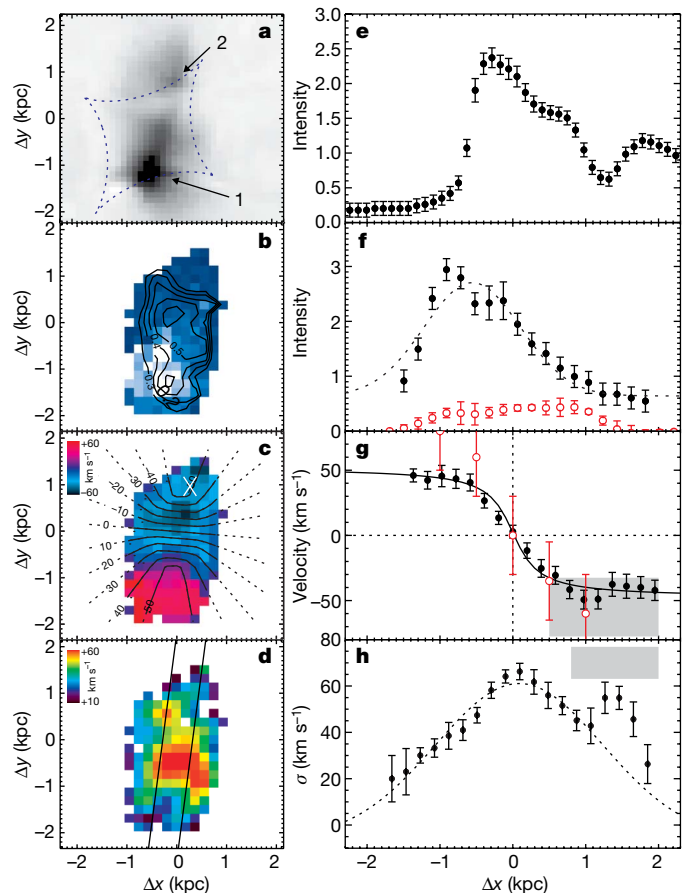


Figure 2 | The internal dynamics of J2135–0102, reconstructed in the source plane with the use of the well-constrained lens model¹². Regions with the largest and smallest magnifications have source-plane resolutions from 0.02 arcsec to 0.04 arcsec, corresponding to physical scales of ~ 150 –300 pc. For clarity we have resampled to a single-pixel scale of 0.02 arcsec. Each pixel therefore represents a nearly independent measurement. **a**, Reconstructed HST V₆₀₆ image of J2135–0102 showing the morphology of the rest-frame ultraviolet stellar continuum. Two major components within the galaxy are labelled 1 and 2. The lens model caustic is overlaid (blue dashed line); regions close to the caustic have the highest magnification. **b**, Distribution of [O III] 5,007 Å emission (blue image) compared with that from H β (contours), which traces the star formation rate with contours varying linearly from 0.2 to 0.6 $M_\odot \text{ yr}^{-1}$ per pixel. **c**, Velocity field derived from the [O III] emission. Overlaid contours indicate the best-fit disk model (see Supplementary Information). X marks the centre of the cold molecular gas¹⁶. **d**, Distribution of the rest-frame [O III] emission linewidth (standard deviation) and corresponding 1σ Poisson uncertainties showing a peak of $54 \pm 4 \text{ km s}^{-1}$ near the centre. **e**, V₆₀₆ surface brightness profile and 1σ Poisson uncertainties. The extraction is performed along the long axis of the galaxy (-10° from the Δy axis) with $\Delta x = 0$ corresponding to a location of (0, -0.2) on the left-hand side. **f**, Spatial profiles and corresponding 1σ Poisson uncertainties of [O III] (black points) and H β (red points). **g**, Major-axis rotation curve for [O III] (black points) and corresponding 1σ uncertainties verified with lower significance using H β (red points) with a best-fitting disk model overlaid (solid line). **h**, Profile of the [O III] emission line standard deviation and corresponding 1σ uncertainties showing a maximum of $\sigma = 54 \pm 4 \text{ km s}^{-1}$. Shaded regions in **g** and **h** show the spatial and kinematic location of the molecular CO emission¹⁶ and the minor influence of component 2 on the overall dynamics.

Supplementary Information) to measure the R_{23} abundance index^{18,19}, finding a metallicity of ~ 0.9 solar metallicity for both components.

These key findings allow us to determine the nature of J2135–0102. The interplay between the dynamics, star formation and mass of the gas reservoir in high-redshift star-forming galaxies is one of the primary drivers of science for the next-generation facilities such as the Atacama Large Millimeter Array (ALMA). Because J2135–0102 is currently the only LBG for which both high-resolution kinematic and cold-gas properties are available, we are uniquely able to constrain the gas mass dynamically and compare this with that implied by local calibrations, thereby testing whether star formation occurs in a mode similar to that seen in local spirals or appropriate to that in more extreme systems. The ratio α between the CO luminosity and the molecular gas mass has been used to distinguish the mode of star formation²⁰. For local spirals $\alpha \approx 5$, which is consistent with star formation occurring in discrete giant molecular clouds with 10^5 – $10^6 M_\odot$. In luminous infrared galaxies, where star formation occurs more vigorously in a single reservoir with $10^9 M_\odot$, α is close to unity²¹. α has yet to be measured at high redshift for all except the most extreme populations²². By requiring that the gas mass be less than the dynamical mass in J2135–0102 we find that $\alpha < 0.8$. This implies that the gas lies in an extensive intercloud medium rather than in discrete, less massive, gravitationally bound clouds²¹. The lower conversion factor identified here suggests that the gas masses of rest-ultraviolet selected galaxies at $z \approx 2$ – 3 may be lower than previously thought²³.

Received 15 January; accepted 21 July 2008.

- Genzel, R. *et al.* The rapid formation of a large rotating disk galaxy three billion years after the Big Bang. *Nature* **442**, 786–789 (2006).
- Forster-Schreiber, N. M. *et al.* SINFONI integral field spectroscopy of $z \sim 2$ UV-selected galaxies. *Astrophys. J.* **645**, 1062–1075 (2006).
- Law, D. *et al.* Integral field spectroscopy of high redshift star-forming galaxies with laser-guide star adaptive optics. *Astrophys. J.* **669**, 929–946 (2007).
- Nesvadba, N. P. H. *et al.* Lyman break galaxies under a microscope: the small-scale dynamics and mass of an arc in the cluster 1E 0657–56. *Astrophys. J.* **650**, 661–668 (2006).
- Ellis, R. S., Santos, M. R., Kneib, J.-P. & Kuijken, K. A faint star-forming system viewed through the lensing cluster Abell 2218: first light at $z = 5.6$? *Astrophys. J.* **560**, 119–122 (2001).
- Kneib, J.-P., Ellis, R. S., Santos, M. R. & Richard, J. A probable $z \sim 7$ galaxy strongly lensed by the rich cluster A2218: exploring the dark ages. *Astrophys. J.* **607**, 697–703 (2004).
- Swinbank, M. *et al.* Resolved spectroscopy of a gravitationally lensed L* Lyman-break galaxy at $z \sim 5$. *Mon. Not. R. Astron. Soc.* **376**, 479–491 (2007).
- Smail, I. *et al.* A very bright highly magnified Lyman break galaxy at $z = 3.07$. *Astrophys. J.* **654**, L33–L36 (2007).
- Dye, S., Smail, I., Swinbank, A. M., Ebeling, H. & Edge, A. C. Separation of the visible and dark matter in the Einstein ring LBG J213512.73–010143. *Mon. Not. R. Astron. Soc.* **379**, 308–316 (2007).
- Larkin, J. *et al.* OSIRIS: a diffraction limited integral field spectrograph for Keck. *N. Astron. Rev.* **50**, 362–364 (2006).
- Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., Broadhurst, T. J. & Franx, M. Galaxy size evolution at high redshift and surface brightness selection effects. *Astrophys. J.* **611**, L1–L4 (2004).
- Abraham, R. *et al.* The star formation history of the Hubble sequence: spatially resolved colour distributions of intermediate-redshift galaxies in the Hubble Deep Field. *Mon. Not. R. Astron. Soc.* **279**, L47–L52 (1996).
- Coppin, K. E. K. *et al.* A detailed study of gas and star formation in a highly magnified Lyman break galaxy at $z = 3.07$. *Astrophys. J.* **665**, 936–943 (2007).
- Kennicutt, R. C. Star formation along the Hubble Sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–232 (1998).
- Erb, D. *et al.* α observations of a large sample of galaxies at $z \sim 2$. *Astrophys. J.* **647**, 128–139 (2006).
- Shapley, A. *et al.* The rest-frame optical properties of $z \sim 3$ galaxies. *Astrophys. J.* **562**, 95–123 (2001).
- Heckman, T. M. Galactic superwinds circa 2001. *ASP Conference Proc.* **254**, 292–304 (2002).
- Pettini, M. & Pagel, B. [O III]/[NII] as an abundance indicator at high redshift. *Mon. Not. R. Astron. Soc.* **348**, L59–L63 (2004).
- Pilyugin, L. S. & Thuan, T. X. Oxygen abundance determination in H II regions: the strong line intensities–abundance calibration revisited. *Astrophys. J.* **631**, 231–243 (2005).
- Solomon, P. M. & Vanden Bout, P. A. Molecular gas at high redshift. *Annu. Rev. Astron. Astrophys.* **43**, 677–725 (2005).
- Solomon, P. M., Downes, D., Radford, S. J. E. & Barrett, J. W. The molecular interstellar medium in ultraluminous infrared galaxies. *Astrophys. J.* **478**, 144–161 (1997).
- Tacconi, L. *et al.* Submillimeter galaxies at $z \sim 2$: evidence for major mergers and constraints on lifetimes, IMF, and CO–H₂ conversion factor. *Astrophys. J.* **680**, 246–262 (2008).
- Erb, D. K. *et al.* The stellar, gas, and dynamical masses of star-forming galaxies at $z \sim 2$. *Astrophys. J.* **646**, 107–132 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Lyke for assistance with the Keck observations, and R. Bower, K. Coppin, M. Lehnert, R. Genzel, D. Erb, D. Law, A. Shapley, A. Jenkins, P. Salucci and T. Theuns for discussions. The OSIRIS data were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership between the California Institute of Technology, the University of California and NASA. The observatory was made possible by the financial support of the W. M. Keck Foundation. The SINFONI data are based on observations made with the ESO Telescopes at the Paranal Observatories. A.M.S. acknowledges financial support from STFC, and I.R.S. and R.S.E. acknowledge financial support from the Royal Society.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.P.S. (dps@astro.caltech.edu).

LETTERS

Observation of the spin Seebeck effect

K. Uchida¹, S. Takahashi^{2,3}, K. Harii¹, J. Ieda^{2,3}, W. Koshibae⁴, K. Ando¹, S. Maekawa^{2,3} & E. Saitoh^{1,5}

The generation of electric voltage by placing a conductor in a temperature gradient is called the Seebeck effect^{1,2}. Its efficiency is represented by the Seebeck coefficient, S , which is defined as the ratio of the generated electric voltage to the temperature difference, and is determined by the scattering rate and the density of the conduction electrons. The effect can be exploited, for example, in thermal electric-power generators and for temperature sensing, by connecting two conductors with different Seebeck coefficients, a device called a thermocouple^{1,2}. Here we report the observation of the thermal generation of driving power, or voltage, for electron spin: the spin Seebeck effect. Using a recently developed spin-detection technique that involves the spin Hall effect^{3–14}, we measure the spin voltage generated from a temperature gradient in a metallic magnet. This thermally induced spin voltage persists even at distances far from the sample ends, and spins can be extracted from every position on the magnet simply by attaching a metal. The spin Seebeck effect observed here is directly applicable to the production of spin-voltage generators, which are crucial for driving spintronic^{15–18} devices. The spin Seebeck effect allows us to pass a pure spin current¹⁹, a flow of electron spins without electric currents, over a long distance. These innovative capabilities will invigorate spintronics research.

The spin Seebeck effect refers to the generation of spin ‘voltage’ as a result of a temperature gradient. We define spin voltage as the spin-current potential, which is represented by $\mu_{\uparrow} - \mu_{\downarrow}$, where μ_{\uparrow} and μ_{\downarrow} respectively denote the electrochemical potentials for spin-up and spin-down electrons^{14,20}; a gradient in the spin voltage drives a spin current. In a metallic magnet, spin-up and spin-down conduction electrons notably have different scattering rates and densities¹⁶, and thus have different Seebeck coefficients, as if two conductors with different S values were inherently present in one magnet (Fig. 1b). When a metallic magnet is subjected to a temperature gradient, therefore, it should generate different driving powers of electrons in different spin channels along the temperature gradient^{21–24}. This is the proposed scenario for the spin Seebeck effect: in the spin sector, a magnet works in the same way as a thermocouple (Fig. 1a, b). This driving power of electrons generates differing amounts of flow in the two spin channels, that is, a spin current. To be more specific, we consider a magnet with uniform magnetization subject to a uniform temperature gradient. In this case, and when the length of the magnet along the temperature gradient is much greater than the spin-diffusion length²⁵ of the magnet, the above scenario and thermodynamic arguments²⁶ predict the spatial distribution of μ_{\uparrow} and μ_{\downarrow} along the temperature gradient shown in Fig. 2a.

However, this spin Seebeck effect remains to be observed. The recently discovered inverse-spin-Hall effect^{10–14} (ISHE) is a powerful tool for detecting spin voltage, and we use it to investigate the spin Seebeck effect in the present study. The ISHE converts a spin current into an electromotive force E_{SHE} by means of spin–orbit scattering. A spin current carries a spin-polarization vector σ along a spatial

direction J_s . The relation between E_{SHE} , J_s and σ is given by the following vector product:

$$E_{\text{SHE}} = D_{\text{ISHE}} J_s \times \sigma \quad (1)$$

The ISHE efficiency D_{ISHE} is enhanced in noble metals, such as Pt. By measuring E_{SHE} , the ISHE can be used to detect a spin current^{10–14}.

Figure 2b shows an illustration of the sample system used in the present study. The sample consists of a 20-nm-thick, soft ferromagnetic $\text{Ni}_{81}\text{Fe}_{19}$ film with a Pt wire attached to one end. The $\text{Ni}_{81}\text{Fe}_{19}$ layer was deposited on a sapphire substrate by electron-beam evaporation in a high vacuum, and the Pt layer was then sputtered in an Ar atmosphere. Immediately before the sputtering, the surface of the $\text{Ni}_{81}\text{Fe}_{19}$ layer was cleaned by Ar ion etching. The length, the width and the thickness of the Pt wire are respectively $L_{\text{Pt}} = 4$ mm, 100 μm , and $d_{\text{Pt}} = 10$ nm. We apply an in-plane external magnetic field, H , along the x direction (Fig. 2d), except when collecting a set of angle-dependent data (discussed below). The coercive force, H_C , of the $\text{Ni}_{81}\text{Fe}_{19}$ layer is around 15 Oe at 300 K, and the magnetization is aligned along the external magnetic field direction when $|H| > H_C$. A temperature gradient ∇T is applied along the x direction by generating a temperature difference ΔT between the ends of the layer. Owing to the direction of the temperature gradient (parallel or antiparallel to

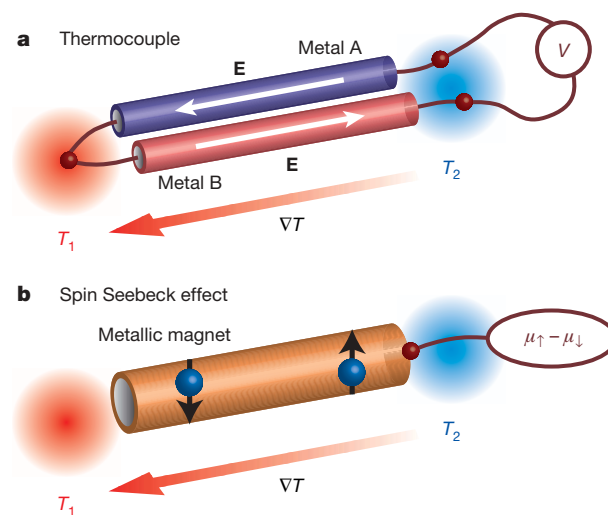


Figure 1 | The spin Seebeck effect. **a**, Illustration of a thermocouple. A thermocouple consists of two conductors (metals A and B) connected to each other. They have different Seebeck coefficients and, thus, the voltage V between the output terminals is proportional to the temperature difference $T_1 - T_2$ between the ends of the couple. **b**, Illustration of the spin Seebeck effect. In a metallic magnet, spin-up (\uparrow) and spin-down (\downarrow) conduction electrons have different Seebeck coefficients. When a temperature gradient is applied, a spin voltage $\mu_{\uparrow} - \mu_{\downarrow}$ proportional to the temperature difference appears; a magnet functions just like a thermocouple, but in the spin sector.

¹Department of Applied Physics and Physico-Informatics, Keio University, Yokohama 223-8522, Japan. ²Institute for Materials Research, Tohoku University, Sendai 980-8577, Japan. ³CREST, Japan Science and Technology Agency, Sanbancho, Tokyo 102-0075, Japan. ⁴Cross-Correlated Materials Research Group, RIKEN, Wako, Saitama 351-0198, Japan. ⁵PRESTO, Japan Science and Technology Agency, Sanbancho, Tokyo 102-0075, Japan.

H), Nernst effects²⁶ are suppressed. The distance between the ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer along the temperature gradient is $L = 6$ mm, which is much greater than the spin-diffusion length of $\text{Ni}_{81}\text{Fe}_{19}$ ²⁵.

The mechanism of the measurement is as follows. In this set-up, the spin voltage, $\mu_{\uparrow} - \mu_{\downarrow}$, which is thermally generated in the $\text{Ni}_{81}\text{Fe}_{19}$ layer, induces a spin current in the Pt layer through the $\text{Ni}_{81}\text{Fe}_{19}$ –Pt interface (Figs 2d and 3a). In the Pt layer, $|\mu_{\uparrow} - \mu_{\downarrow}|$ decreases along the z direction. Because the spin polarization, σ , of this spin current lies along the magnetization direction, when $|H| > H_C$ the spin current

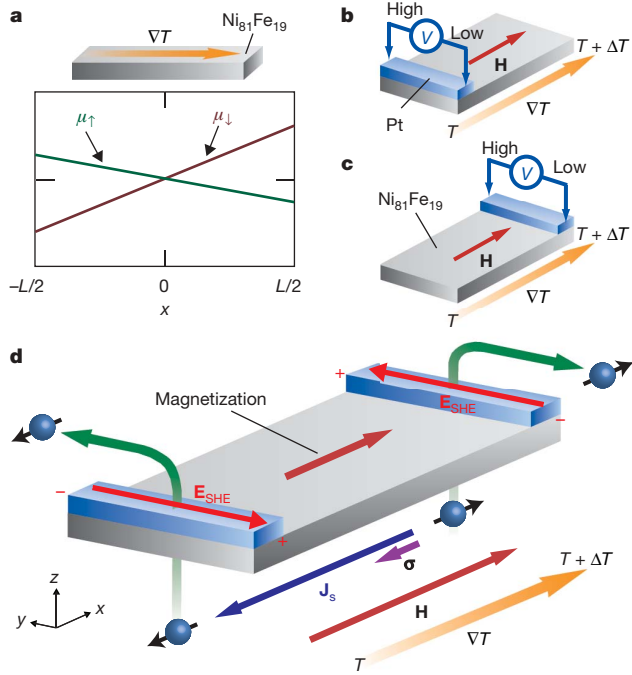


Figure 2 | Experimental set-up. **a**, Illustration of the calculated distributions of the electrochemical potentials μ_{\uparrow} and μ_{\downarrow} for spin-up and spin-down electrons in a metallic magnet. A temperature gradient ∇T is applied along the x direction. We write μ_{\uparrow} and μ_{\downarrow} respectively as $\mu_{\uparrow} = \mu_{\uparrow}^{\text{c}} - e\phi$ and $\mu_{\downarrow} = \mu_{\downarrow}^{\text{c}} - e\phi$, where e is the elementary charge, ϕ is the electrostatic potential, and $\mu_{\uparrow}^{\text{c}}$ and $\mu_{\downarrow}^{\text{c}}$ are the spin-dependent chemical potentials, which depend on the temperature T and on the densities n_{\uparrow} and n_{\downarrow} . Thus, the gradient of μ_{\uparrow} can be expanded as $\nabla\mu_{\uparrow} = (\partial\mu_{\uparrow}^{\text{c}}/\partial n_{\uparrow})\nabla n_{\uparrow} + (\partial\mu_{\uparrow}^{\text{c}}/\partial T)\nabla T - e\nabla\phi$ (and that of μ_{\downarrow} similarly). The first term on the right-hand side reflects the particle-density (n_{\uparrow}) modulation (accumulation term) and the second term the entropy contribution; $\partial\mu_{\uparrow}^{\text{c}}/\partial T$ reflects the spin-subband entropy. Because, in comparison with our sample size ($L = 6$ mm), ∇n_{\uparrow} and ∇n_{\downarrow} decay within a spin-diffusion length²⁵ (~ 5 nm in NiFe) of the ends, the accumulation terms have a minor role and the difference between $\nabla\mu_{\uparrow}$ and $\nabla\mu_{\downarrow}$ is given by $\nabla(\mu_{\uparrow} - \mu_{\downarrow}) = eS_S\nabla T$, where $S_S = (1/e)[\partial\mu_{\uparrow}^{\text{c}}/\partial T - \partial\mu_{\downarrow}^{\text{c}}/\partial T]$. For a metallic magnet, in general, $\partial\mu_{\uparrow}^{\text{c}}/\partial T \neq \partial\mu_{\downarrow}^{\text{c}}/\partial T$. This implies that a temperature gradient induces a finite gradient of spin voltage, $\mu_{\uparrow} - \mu_{\downarrow}$. Because of the symmetry, $\mu_{\uparrow} - \mu_{\downarrow}$ crosses zero at the centre of the sample. **b, c**, Illustrations of the measurement set-up in the present study. The sample consists of a $\text{Ni}_{81}\text{Fe}_{19}$ film with a Pt wire attached to one end. The surface of the $\text{Ni}_{81}\text{Fe}_{19}$ film is a 24-mm^2 rectangle. The chemical composition was confirmed using fluorescent X-ray microscopy. **H** denotes the external magnetic field vector (with magnitude H). To generate a temperature gradient in the film, the temperature difference between two Cu blocks connected to either end of the sample was controlled in a vacuum, using heaters and thermocouples attached to the blocks. **d**, Illustration of the spin Seebeck effect induced by a temperature gradient ∇T in a $\text{Ni}_{81}\text{Fe}_{19}$ film and the ISHE induced in Pt wires attached to the ends of film. Here \mathbf{J}_s and \mathbf{E}_{SHE} respectively denote the spatial direction of the thermally induced spin current in the $\text{Ni}_{81}\text{Fe}_{19}$ film and the electromotive force generated by the ISHE in the Pt wire. The spin polarization vector, σ , of the spin current lies along the magnetization direction.

induced in the Pt layer generates an electromotive force along the y direction owing to the ISHE in the Pt layer (Fig. 2d and equation (1)). Here we note that the sign of $\mu_{\uparrow} - \mu_{\downarrow}$ reverses between the ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer, as illustrated in Fig. 2a. Therefore, the direction of σ in the Pt layer and the sign of \mathbf{E}_{SHE} , the electromotive force induced by the ISHE, are expected to reverse between the higher- and the lower-temperature ends (Fig. 2d and equation (1)). This distinctive behaviour of the electromotive force is characteristic of the ISHE induced

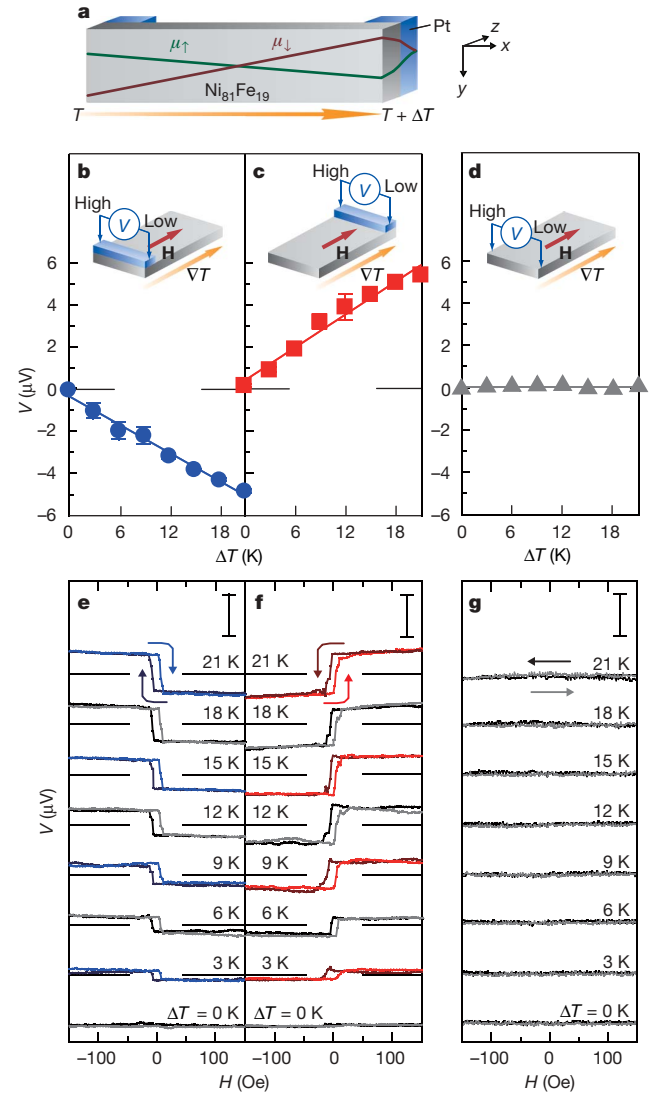


Figure 3 | Measurements of electromotive force. **a**, Illustration of the spin-dependent electrochemical potentials, μ_{\uparrow} and μ_{\downarrow} , induced by a temperature gradient in the $\text{Ni}_{81}\text{Fe}_{19}$ –Pt sample system. ΔT denotes the temperature difference between the ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. **b, c**, ΔT dependence of the electric voltage difference V between the ends of the Pt wire in the $\text{Ni}_{81}\text{Fe}_{19}$ –Pt sample for $H = 100$ Oe when the Pt wire is attached to the lower-temperature (300 K, **b**) and higher-temperature (300 K + ΔT , **c**) ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. For $H = 100$ Oe, the magnetization of the $\text{Ni}_{81}\text{Fe}_{19}$ layer is aligned along the magnetic field direction. The error bars represent the 95% confidence level ($\pm 2\text{s.d.}$). **d**, ΔT dependence of the electric voltage difference V between the sides of a plain $\text{Ni}_{81}\text{Fe}_{19}$ film (no Pt layer) for $H = 100$ Oe, in the direction perpendicular to that of the temperature-gradient. The measurement was performed at the lower-temperature end of the film. The errors are negligibly small. **e, f**, H dependence of V between the ends of the Pt wire in the $\text{Ni}_{81}\text{Fe}_{19}$ –Pt sample for various values of ΔT , measured when the Pt wire is attached to the lower-temperature (**e**) and higher-temperature (**f**) ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. **g**, H dependence of V between the sides of a plain $\text{Ni}_{81}\text{Fe}_{19}$ film, in the direction perpendicular to that of temperature gradient. The measurement was performed at the lower-temperature end of the film. Scale bars in **e, f** and **g**, $10\text{ }\mu\text{V}$.

by the thermally generated spin voltage. Hence, measuring this electromotive force allows us to detect the spin Seebeck effect with high sensitivity. To do this, we measure the electric voltage difference, V , between the ends of the Pt wire of the sample system at 300 K, as illustrated in Fig. 2b, c.

Figure 3b and Fig. 3c show V at $H = 100$ Oe ($> H_C$) as a function of ΔT when the Pt wire is on the lower- and, respectively, higher-temperature ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. Measurements were made using identical samples. The magnitude of V is proportional to ΔT at both ends of the layer. Figure 3d shows that the electric voltage signal disappears in a plain $\text{Ni}_{81}\text{Fe}_{19}$ film to which no Pt wire is attached (see also Fig. 3g). This signal also disappears in a plain Pt film. These observations clearly demonstrate that the observed V signal shown in Fig. 3b, c is not associated with the conventional Seebeck effect or the Nernst effects²⁶ in either an unconnected $\text{Ni}_{81}\text{Fe}_{19}$ film or a Pt layer, but is only generated when the two layers are electrically connected.

Significantly, Fig. 3b, c also shows that the sign of V for finite values of ΔT clearly reverses between the lower- and higher-temperature ends of the layer. This unconventional behaviour of V is the feature predicted for the ISHE induced by the spin Seebeck effect described above (Fig. 2d). This is therefore direct evidence for the operation of the spin Seebeck effect in the $\text{Ni}_{81}\text{Fe}_{19}$ film.

To further support this interpretation, we measure the magnetic field dependence of V in the same system. In Fig. 3e and Fig. 3f, we show V as a function of H for various values of ΔT , measured at the lower- and, respectively, higher-temperature ends of the layer. The external magnetic field is applied along the x direction. Notably, in the low-field range, -15 Oe $< H < 15$ Oe, hysteresis loops appear in all the data. These hysteresis loops correspond to the magnetization hysteresis that arises in the $\text{Ni}_{81}\text{Fe}_{19}$ layer in response to the variation of H . This demonstrates that the V signal is affected by the magnetization, that is, the spin-polarization direction in the $\text{Ni}_{81}\text{Fe}_{19}$ layer. When $|H| > H_C$, the sign of V at each end of the layer can be reversed by reversing H . This sign reversal is consistent with the prediction of the ISHE induced by the spin Seebeck effect, through equation (1) (we note that σ is reversed by a magnetization reversal).

Figure 4a shows the voltage difference, V , measured with respect to θ , the angle the in-plane magnetic field makes with the x direction.

We fix the temperature difference, ΔT , and the magnetic field strength, H , at 10 K and 100 Oe. Because $H > H_C$, the magnetization in the $\text{Ni}_{81}\text{Fe}_{19}$ layer and σ are aligned along the external magnetic field direction. At both the lower- and the higher-temperature ends of the layer, V varies with θ in a sinusoidal pattern and vanishes when $\theta = 90^\circ$, a situation consistent with the prediction of the ISHE described in equation (1). The sign of V reverses between the lower- and higher-temperature ends for each θ value, a result which also supports the aforementioned prediction of the ISHE induced by the spin Seebeck effect (Fig. 2d).

The spin voltage induced at the higher-temperature end of the $\text{Ni}_{81}\text{Fe}_{19}$ layer with temperature difference ΔT is $(\mu_\uparrow - \mu_\downarrow) = eS_S\Delta T/2$, where S_S is the spin Seebeck coefficient. This spin voltage injects spins into the Pt layer and generates the electric voltage $V \approx \theta_{\text{Pt}}\eta_{\text{NiFe-Pt}}(L_{\text{Pt}}/d_{\text{Pt}})S_S\Delta T/2$ owing to the ISHE, where θ_{Pt} is the spin Hall angle^{11–14} of Pt and $\eta_{\text{NiFe-Pt}}$ is the spin-injection efficiency¹⁴. It is important to note that V is greatly amplified by the very large aspect ratio, $L_{\text{Pt}}/d_{\text{Pt}} \approx 4 \times 10^5$, of the Pt layer, despite the small temperature gradient in the $\text{Ni}_{81}\text{Fe}_{19}$ layer. For $V/\Delta T = 0.25 \mu\text{V K}^{-1}$, $\eta_{\text{NiFe-Pt}} \approx 0.2$ and $\theta_{\text{Pt}} = 0.0037$ (ref. 12), we estimate the spin Seebeck coefficient for $\text{Ni}_{81}\text{Fe}_{19}$ to be $S_S = -2 \text{ nV K}^{-1}$ at 300 K, which is different from the conventional Seebeck coefficient of $-20 \mu\text{V K}^{-1}$ measured for the present $\text{Ni}_{81}\text{Fe}_{19}$ film. This implies that the spin relaxation decreases spin accumulation in this large-area film.

Up to now, we have discussed the spin voltage at the ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. It is also interesting to enquire whether the spin voltage obeys the calculated spatial variation illustrated in Fig. 2a far from the ends. To investigate the spatial distribution of the spin voltage, we measure V using Pt wires placed at different positions on the $\text{Ni}_{81}\text{Fe}_{19}$ films. In each case, the Pt wires are placed perpendicular to the x direction (Fig. 4b). During the measurement, an external magnetic field of 100 Oe ($> H_C$) is applied along the x direction. In Fig. 4b, we show V as a function of x_p , the position of the Pt wire along the temperature gradient from the centre of the $\text{Ni}_{81}\text{Fe}_{19}$ layer, for various values of ΔT . With ΔT applied, V clearly increases for $x_p > 0$ and decreases for $x_p < 0$. For each value of ΔT , V varies almost linearly with x_p . The calculation shown in Fig. 2a reproduces these experimental results very well, on the basis that V is proportional to

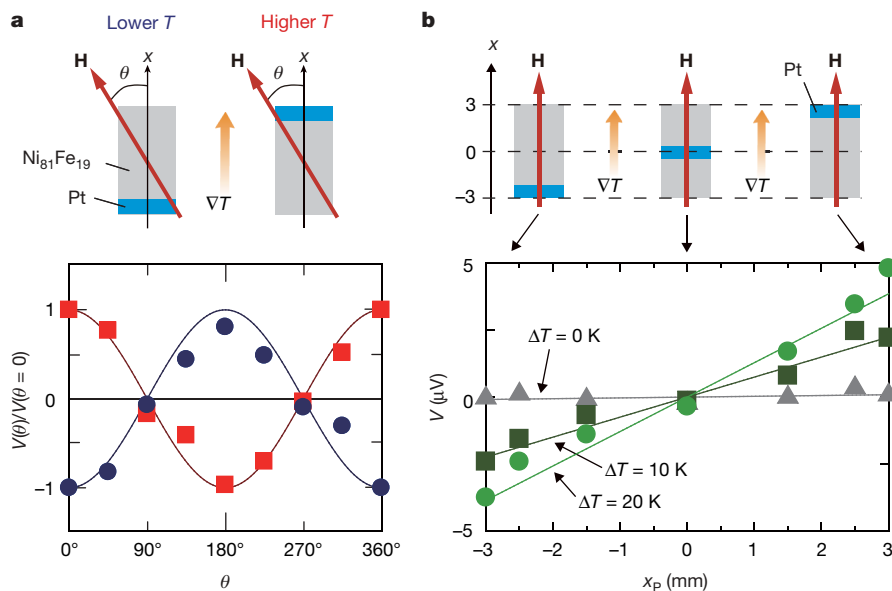


Figure 4 | Dependence on magnetic field angle and position of Pt wire. **a**, Dependence on magnetic field angle of the electric voltage difference V between the ends of the Pt wire in the $\text{Ni}_{81}\text{Fe}_{19}$ -Pt sample, for $H = 100$ Oe and $\Delta T = 10$ K, when the Pt wire is attached to the lower-temperature (blue circles) and higher-temperature (red squares) ends of the $\text{Ni}_{81}\text{Fe}_{19}$ layer. An

in-plane magnetic field is applied at an angle θ to the x direction. **b**, Dependence of V on x_p , the position of the Pt wire, measured using four separate samples and seven different values of x_p on $\text{Ni}_{81}\text{Fe}_{19}$ films. In all the samples, the Pt wires are placed perpendicular to the x direction and an external magnetic field of 100 Oe ($> H_C$) is applied along the x direction.

$\mu_{\uparrow} - \mu_{\downarrow}$ at each position on the $\text{Ni}_{81}\text{Fe}_{19}$ layer. This experiment yields the result that, when using the spin Seebeck effect, spins can be extracted from every position on the $\text{Ni}_{81}\text{Fe}_{19}$ surface simply by attaching a metal contact.

The spin Seebeck effect can be applied directly to the construction of thermal spin generators for driving spintronic devices, thereby opening the door to thermo-spintronics. The present study has also shown that, in combination with the ISHE, the spin Seebeck effect allows us to make effective thermoelectric generators and temperature-gradient sensors, as the electric voltage generated can be amplified simply by increasing the length, L_{Pt} , of the Pt wire. The spin Seebeck effect generates a pure spin current over a long distance in a magnet, without electric currents. This is in striking contrast to any of the other known methods for the injection of pure spin current^{10–14,27,28}, where spin current and spin voltage disappear over distances greater than the spin-diffusion length²⁵ (for example ~ 500 nm in Cu and ~ 5 nm in Pt). This length scale severely limits experiments on pure spin currents. The long-distance nature of the spin Seebeck effect observed here immediately suggests a way of applying a variety of experimental tools to spin currents. We anticipate that the spin Seebeck effect will transform basic research on spin currents and lead to dramatic advances in spintronic devices.

Received 6 May; accepted 4 August 2008.

- Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* 253–258 (Saunders College, 1976).
- Maekawa, S. et al. *Physics of Transition Metal Oxides* 323–331 (Springer, 2004).
- Dyakonov, M. I. & Perel, V. I. Current-induced spin orientation of electrons in semiconductors. *Phys. Lett. A* **35**, 459–460 (1971).
- Hirsch, J. E. Spin Hall effect. *Phys. Rev. Lett.* **83**, 1834–1837 (1999).
- Takahashi, S. & Maekawa, S. Hall effect induced by a spin-polarized current in superconductors. *Phys. Rev. Lett.* **88**, 116601 (2002).
- Murakami, S., Nagaosa, N. & Zhang, S.-C. Dissipationless quantum spin current at room temperature. *Science* **301**, 1348–1351 (2003).
- Sinova, J. et al. Universal intrinsic spin Hall effect. *Phys. Rev. Lett.* **92**, 126603 (2004).
- Kato, Y. K., Myers, R. C., Gossard, A. C. & Awschalom, D. D. Observation of the spin Hall effect in semiconductors. *Science* **306**, 1910–1913 (2004).
- Wunderlich, J., Kaestner, B., Sinova, J. & Jungwirth, T. Experimental observation of the spin-Hall effect in a two-dimensional spin-orbit coupled semiconductor system. *Phys. Rev. Lett.* **94**, 047204 (2005).
- Saitoh, E., Ueda, M., Miyajima, H. & Tataru, G. Conversion of spin current into charge current at room temperature: Inverse spin-Hall effect. *Appl. Phys. Lett.* **88**, 182509 (2006).
- Valenzuela, S. O. & Tinkham, M. Direct electronic measurement of the spin Hall effect. *Nature* **442**, 176–179 (2006).
- Kimura, T., Otani, Y., Sato, T., Takahashi, S. & Maekawa, S. Room-temperature reversible spin Hall effect. *Phys. Rev. Lett.* **98**, 156601 (2007).
- Seki, T. et al. Giant spin Hall effect in perpendicularly spin-polarized FePt/Au devices. *Nature Mater.* **7**, 125–129 (2008).
- Takahashi, S. & Maekawa, S. Spin current in metals and superconductors. *J. Phys. Soc. Jpn* **77**, 031009 (2008).
- Wolf, S. A. et al. Spintronics: a spin-based electronics vision for the future. *Science* **294**, 1488–1495 (2001).
- Žutić, I., Fabian, J. & Das Sarma, S. Spintronics: fundamentals and applications. *Rev. Mod. Phys.* **76**, 323–410 (2004).
- Maekawa, S. (ed.) *Concepts in Spin Electronics* (Oxford Univ. Press, 2006).
- Chappert, C., Fert, A. & Van Dau, F. N. The emergence of spin electronics in data storage. *Nature Mater.* **6**, 813–823 (2007).
- Slonczewski, J. C. Conductance and exchange coupling of two ferromagnets separated by a tunneling barrier. *Phys. Rev. B* **39**, 6995–7002 (1989).
- Valet, T. & Fert, A. Theory of the perpendicular magnetoresistance in magnetic multilayers. *Phys. Rev. B* **48**, 7099–7113 (1993).
- Cadeville, M. C. & Roussel, J. Thermoelectric power and electronic structure of dilute alloys of nickel and cobalt with d transition elements. *J. Phys. F* **1**, 686–710 (1971).
- Gravier, L., Serrano-Guisan, S., Reuse, F. & Ansermet, J.-P. Thermodynamic description of heat and spin transport in magnetic nanostructures. *Phys. Rev. B* **73**, 024419 (2006).
- Tsyplatyev, O., Kashuba, O. & Fal'ko, V. I. Thermally excited spin current and giant magnetothermopower in metals with embedded ferromagnetic nanoclusters. *Phys. Rev. B* **74**, 132403 (2006).
- Hatami, M., Bauer, G. E. W., Zhang, Q.-F. & Kelly, P. J. Thermal spin-transfer torque in magnetoelectronic devices. *Phys. Rev. Lett.* **99**, 066603 (2007).
- Bass, J. & Pratt, W. P. Jr. Spin-diffusion lengths in metals and alloys, and spin-flipping at metal/metal interfaces: an experimentalist's critical review. *J. Phys. Condens. Matter* **19**, 183201 (2007).
- Callen, H. B. *Thermodynamics* Ch. 17 (Wiley, 1960).
- Hirohata, A., Xu, Y. B., Guertler, C. M., Bland, J. A. C. & Holmes, S. N. Spin-polarized electron transport in ferromagnet/semiconductor hybrid structures induced by photon excitation. *Phys. Rev. B* **63**, 104425 (2001).
- Jedema, F. J., Filip, A. T. & van Wees, B. J. Electrical spin injection and accumulation at room temperature in an all-metal mesoscopic spin valve. *Nature* **410**, 345–348 (2001).

Acknowledgements The authors thank Y. Suzuki, S. E. Barnes, Y. Fujitani, G. Tataru, K. M. Itoh, H. Kuwahara and M. Matoba for discussions. This work was supported by a Grant-in-Aid for Scientific Research in Priority Area 'Creation and control of spin current' (19048028) from MEXT, Japan, a Grant-in-Aid for Encouragement of Young Scientists (A) from MEXT, Japan, the global COE for the 'High-level global cooperation for leading-edge platform on access spaces' from MEXT, Japan, a Strategic Information and Communications R&D Promotion Programme from MIC, Japan, and the Next Generation Supercomputing Project of Nanoscience Program from IMS, Japan.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.S. (saitoheiji@appi.keio.ac.jp).

LETTERS

High-temperature interface superconductivity between metallic and insulating copper oxides

A. Gozar¹, G. Logvenov¹, L. Fitting Kourkoutis², A. T. Bollinger¹, L. A. Giannuzzi³, D. A. Muller² & I. Bozovic¹

The realization of high-transition-temperature (high- T_c) superconductivity confined to nanometre-sized interfaces has been a long-standing goal because of potential applications^{1,2} and the opportunity to study quantum phenomena in reduced dimensions^{3,4}. This has been, however, a challenging target: in conventional metals, the high electron density restricts interface effects (such as carrier depletion or accumulation) to a region much narrower than the coherence length, which is the scale necessary for superconductivity to occur. By contrast, in copper oxides the carrier density is low whereas T_c is high and the coherence length very short, which provides an opportunity—but at a price: the interface must be atomically perfect. Here we report superconductivity in bilayers consisting of an insulator (La_2CuO_4) and a metal ($\text{La}_{1.55}\text{Sr}_{0.45}\text{CuO}_4$), neither of which is superconducting in isolation. In these bilayers, T_c is either ~ 15 K or ~ 30 K, depending on the layering sequence. This highly robust phenomenon is confined within 2–3 nm of the interface. If such a bilayer is exposed to ozone, T_c exceeds 50 K, and this enhanced superconductivity is also shown to originate from an interface layer about 1–2 unit cells thick. Enhancement of T_c in bilayer systems was observed previously⁵ but the essential role of the interface was not recognized at the time.

Typical approaches to the realization of quasi-two-dimensional superconducting sheets rely on fabrication of an ‘ultrathin’ layer of a known superconductor^{1,2}. Another route is to use hetero-interfaces. Superconductivity in the 0.2–6 K range was reported at the interface between two oxide insulators⁶ and in superlattices where one⁷ or both⁸ components are semiconductors. The $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ (LSCO) family is particularly attractive because these materials are solid solutions that can be doped over a broad range⁹.

In our experiment, we have synthesized a large number (over 200) of single-phase, bilayer and trilayer films with insulating (I), metallic (M) and superconducting (S) blocks in all combinations and of varying layer thickness (for notation, see Fig. 1 legend). The films were grown in a unique atomic-layer-by-layer molecular beam epitaxy system¹⁰ that incorporates *in situ* state-of-the-art surface science tools, such as time-of-flight ion scattering and recoil spectroscopy, and reflection high-energy electron diffraction. It enables synthesis of atomically smooth films as well as multilayers with perfect interfaces^{5,11–13}. Typical surface roughnesses determined from atomic force microscopy data are 0.2–0.5 nm, less than one unit cell, which in LSCO is 1.3 nm. Atomic-layer-by-layer molecular beam epitaxy provides for digital control of layer thickness, which we measure by counting the number of unit cells. Maintaining atomic-scale smoothness and digital layer-by-layer growth are both crucial for the results we discuss in the following.

The interface between the metallic and insulating materials is superconducting with high T_c (Fig. 1), and the deposition sequence

matters. M–S bilayers show the highest T_c , ~ 50 K. In contrast, in single-phase LSCO films that we have grown under the same conditions, the highest T_c is about 40 K, similar to what is seen in bulk single crystals (ref. 9 and Supplementary Fig. 1). Hence, in M–S bilayers we see a large (up to 25%) relative T_c enhancement. T_c values around 50 K were observed previously in some LSCO films^{14,15} and

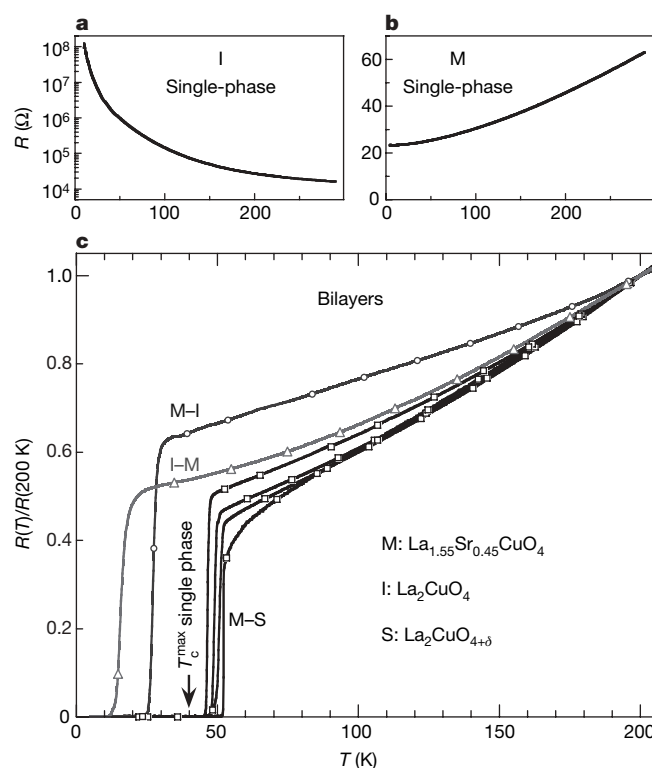


Figure 1 | The dependence of resistance on temperature for single-phase and bilayer films. Notation used in text and figures is as follows: I is La_2CuO_4 , vacuum-annealed and insulating; S is $\text{La}_2\text{CuO}_{4+\delta}$, oxygen-doped by annealing in ozone and superconducting; M is $\text{La}_{1.55}\text{Sr}_{0.45}\text{CuO}_4$, overdoped and metallic but not superconducting; R , resistance; T , temperature. For bilayers, the first letter always denotes the layer next to the LaSrAlO_4 substrate. **a**, **b**, $R(T)$ for single-phase layers of I (**a**; note the log scale) and M (**b**). **c**, $R(T)$ normalized to $T = 200$ K for various bilayers. Typical values for T_c at the mid-point of the resistive transitions are $T_c \approx 15$ K in I–M and $T_c \approx 30$ K in M–I structures. In M–S bilayers (four samples shown), $T_c \approx 50$ K. In a few hundred single-phase films (doped by either oxygen or Sr) grown under the same conditions, T_c never exceeded 40 K, the value marked by the arrow (Supplementary Fig. 1). The interface superconductivity is reproducible and stable in air on the scale of months in contrast to single-phase S films.

¹Brookhaven National Laboratory, Upton, New York 11973-5000, USA. ²School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA. ³FEI Company, Hillsboro, Oregon 97124, USA.

LSCO–LCO bilayers⁵, but the locus of the highest T_c has not been investigated. We show below that in our M–I films, enhanced superconductivity originates from and is restricted to an interfacial layer 1–2 unit cells thick. In retrospect, one would suppose that at least the bilayer result⁵ was also an interface effect, a proposition that we confirmed, as discussed below.

To directly determine the length scale associated with interface superconductivity, we synthesized a series of M–I and I–M structures with thick bottom layers (≥ 30 unit cells) while the thickness of the top layer was increased digitally, one-half a unit cell at a time (Fig. 2). The transport data show that the plateau values for superconductivity are reached after the thickness of the top layer is ≥ 2 unit cells, a value that sets the length scale for this interface phenomenon.

The T_c enhancement in M–S bilayers triggers the intriguing question as to whether this enhancement is an interface phenomenon, as suggested by several preliminary observations (Supplementary

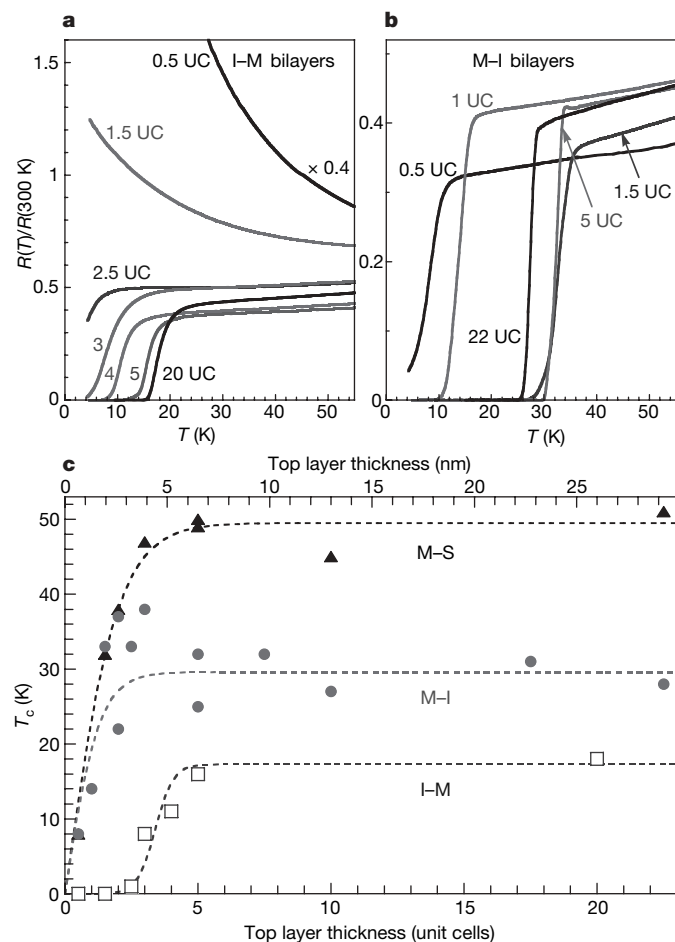


Figure 2 | The dependence on the layer thickness. **a**, Normalized resistance for several I–M bilayers where the thickness of the bottom I layer is fixed at 40 unit cells, that is, 52 nm, and the thickness of the M layer is varied as indicated (UC, unit cell). For an M layer 0.5 unit cells thick the sample is insulating, whereas the 1.5 unit cell structure shows a metallic-to-insulating crossover near $T = 75$ K. Further increase of the thickness raises T_c to a plateau of 15 K. **b**, The same for M–I bilayers with a 40-unit-cell-thick bottom M layer. Traces of superconductivity can be observed even when the bottom M layer is covered by an I layer only 0.5 unit cell thick (0.66 nm). When one unit cell of I covers the surface, the resistive transition is complete and $T_c > 10$ K. On its own, this is a signature of virtually atomically perfect surfaces, given that the resistance measurements were taken with the voltage probes 3 mm apart. **c**, T_c (defined as the midpoint of the resistive transition) as a function of the top layer thickness in M–I, I–M and M–S bilayers. The last are structures obtained by annealing M–I bilayers in an ozone atmosphere, the procedure that turns I films into S but has essentially no effect on M. Dashed lines are guides for the eye.

Information). That this is the case is confirmed by the data on critical current density (j_c) determined from two-coil mutual inductance measurements^{16–18} (Fig. 3). The results indicate that the $T_c \approx 50$ K in M–S structures is in fact confined to a very thin (1–2 unit cells thick) layer near the interface. The observed linear temperature dependence of j_c in S films is expected theoretically in copper oxides for the intrinsic critical current due to vortex–antivortex pair breaking or depinning in homogeneous samples¹⁹, and it is observed experimentally in high-quality films and bulk single crystals of high-temperature superconductors²⁰. In contrast, in M–S samples, one can see a clear break near 40 K that separates two approximately linear regions with very different slopes.

This is what one expects from two superconducting sheets with different thicknesses and transition temperatures, say d_1 , T_{c1} and d_2 , T_{c2} , respectively. The breakdown into two such components (the dashed lines in Fig. 3) provides $T_{c1} \approx 40$ K and $T_{c2} \approx 50$ K. The

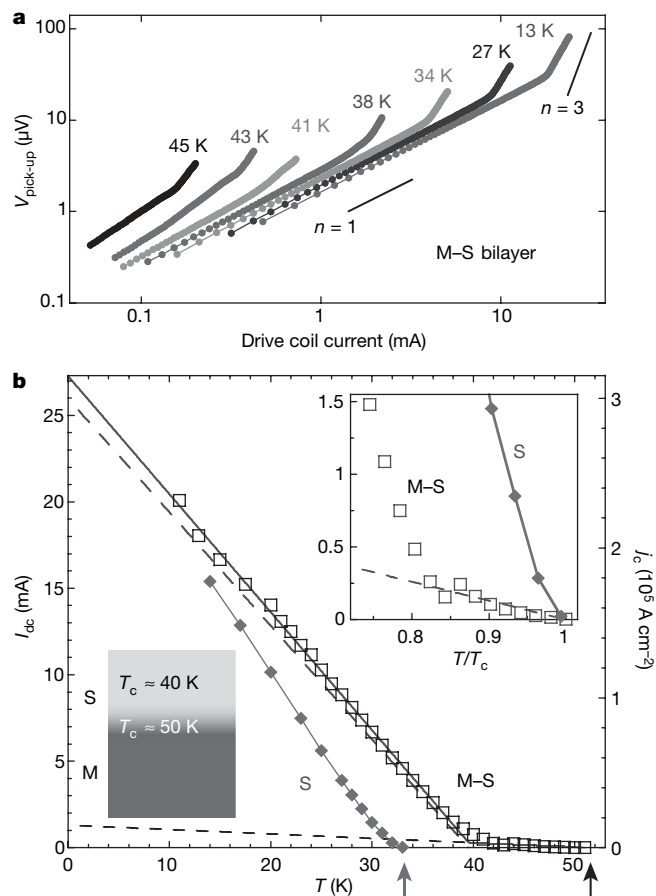


Figure 3 | Nonlinear screening effects in a single-phase S film and an M–S bilayer. **a**, The dependence of the pick-up voltage on the current in the drive coil at several temperatures. At each temperature, a ‘critical’ value of the current in the drive coil, I_{dc} , corresponds to the onset of dissipation in the film, and can be defined as the crossover point between a linear ($n = 1$) and a higher-power law ($n \approx 3$ at temperatures below 40 K) behaviour. In both samples, the S layer is 20 unit cells thick. **b**, The temperature dependence of I_{dc} for an S film (filled diamonds) and an M–S bilayer (open squares). The right scale shows the calculated peak value of the induced screening current density in superconducting films (see also Supplementary Fig. 5). Arrows denote values of T_c : 33.2 K and 51.6 K for the S and M–S samples, respectively. The bilayer data can be well decomposed into two approximately linear contributions (dashed lines), corresponding to bulk and interface parts with $T_c \approx 40$ K and $T_c \approx 50$ K (lower left inset). Top right inset shows the same data in reduced temperature units, T/T_c . The magnitude of the estimated low-temperature critical current of the thin layer is in agreement with the value estimated from mutual inductance and transport measurements in M–I bilayers in which the high-temperature superconductor ($T_c = 30$ K) sheet has a similar thickness.

low-temperature extrapolation of the critical current gives $d_1/d_2 \approx 20$. As the total number of layers deposited was $d_1 + d_2 = 20$ unit cells, one obtains $d_2 \approx 1$ unit cell. This length scale is quantitatively consistent with the independent measurements of resistivity in M–S bilayers as a function of top layer thicknesses (Fig. 2c). We performed similar mutual inductance measurements on the same bilayer sample (which had not deteriorated after seven years) that was studied in ref. 5, in which the bottom layer was optimally doped LSCO, and the results were quite similar to the M–S case; this demonstrates that the previously reported T_c enhancement was also an interface effect.

We now consider the issue of interface structure and the possible impact of cation interdiffusion (Fig. 4). The microstructure of an M–I bilayer and its interfaces was analysed using electron energy-loss spectroscopy in a scanning transmission electron microscope. An upper limit on the amount of chemical interdiffusion at the interfaces is obtained by recording the lanthanum-M_{4,5} edges in the spectra. The root-mean-square interface roughness, as determined by fitting error functions to the La profile, is $\sigma = 0.8 \pm 0.4$ nm at the substrate–M interface and $\sigma = 1.2 \pm 0.4$ nm (~ 1 unit cell) at the M–I interface, which sets an upper limit to any cation intermixing (see also Supplementary Fig. 8).

As an independent test of chemical variations across the interfaces, the changes in the oxygen-K fine structure were analysed using a principal-components analysis. The fraction of the component

corresponding to the M layer is shown in Fig. 4d, which again indicates an interface roughness of less than one unit cell. Either interface was fully described by two components, leaving no significant residual after the fit, suggesting that there is no substantial third, interfacial layer, at least on the scale of the interface roughness. Results obtained by several other surface-sensitive probes, such as reflection high-energy electron diffraction, and time-of-flight ion scattering and recoil spectroscopy, as well as transport on I–M–I heterostructures (Supplementary Figs 2, 3 and 4), support and are consistent with the chemically abrupt interfaces inferred from the scanning transmission electron microscope data. The experiments set an upper limit on possible cation interdiffusion of less than 1 unit cell depth, and make the cation mixing scenario hard to reconcile quantitatively and qualitatively with our observations.

Other possible causes for the high-temperature superconductivity at the interface are electronic reconstruction or oxygen non-stoichiometry. Experimental data show that charge depletion or accumulation is substantial across M–I and I–M interfaces²¹, whereas such charge transfer is negligible when M is replaced by optimally doped LSCO (ref. 15). These findings are consistent with the doping dependence of the chemical potential in LSCO inferred from X-ray photoemission data²². Oxygen vacancies and interstitials are nevertheless additional factors that should be considered: they may account for the asymmetry between M–I and I–M structures, and are essential for the increased T_c and stability of superconductivity in M–S bilayers (Supplementary Information, section E).

A remaining puzzle is the mechanism of relative T_c enhancement in M–S bilayers. It is conceivable that structural aspects, such as disorder, play a crucial role. We may have realized the doping without disorder scenario²³ by introducing carriers via charge transfer and by (ordered) interstitial oxygen pinned near the interface. Another possibility is that the ‘intrinsic’ T_c in LSCO would be even higher were it not for some competing instability, and that this other order parameter is suppressed in bilayers via the long-range strain or electrostatic effects (or both). Finally, an interesting possibility is that pairing and/or coherence of electrons in one layer is enabled or enhanced by interactions originating in the neighbouring layer^{24,25}. Deciphering this problem may open the path to even larger T_c enhancement.

Received 15 June 2007; accepted 25 July 2008.

- Ahn, C. H., Triscone, J.-M. & Mannhart, J. Electric field effect in correlated oxide systems. *Nature* **424**, 1015–1018 (2003).
- Ahn, C. H. *et al.* Electrostatic modulation of superconductivity in ultrathin GdBa₂Cu₃O_{7- δ} . *Science* **284**, 1152–1155 (1999).
- Berezinskii, V. L. Destruction of long-range order in one-dimensional and 2-dimensional systems having a continuous symmetry group 2. Quantum systems. *Sov. Phys. JETP* **34**, 610–616 (1972).
- Kosterlitz, J. M. & Thouless, D. J. Ordering, metastability and phase transitions in two-dimensional systems. *J. Phys. C* **6**, 1181–1203 (1973).
- Bozovic, I., Logvenov, G., Belca, I., Narimbetov, B. & Sveklo, I. Epitaxial strain and superconductivity in La_{2-x}Sr_xCuO₄ thin films. *Phys. Rev. Lett.* **89**, 107001 (2002).
- Reyren, N. *et al.* Superconducting interfaces between insulating oxides. *Science* **317**, 1196–1199 (2007).
- Seguchi, Y., Tsuboi, T. & Suzuki, T. Magnetic-field-enhanced superconductivity in Au/Ge layered films. *J. Phys. Soc. Jpn* **61**, 1875–1878 (1992).
- Fogel, N. Ya. *et al.* Interfacial superconductivity in semiconducting monochalcogenide superlattices. *Phys. Rev. B* **73**, R161306 (2006).
- Kastner, M. A. & Birgeneau, R. J. Magnetic, transport and optical properties of monolayer copper oxides. *Rev. Mod. Phys.* **70**, 897–928 (1998).
- Bozovic, I. Atomic layer engineering of superconducting oxides: Yesterday, today, tomorrow. *IEEE Trans. Appl. Supercond.* **11**, 2686–2695 (2001).
- Bozovic, I., Eckstein, J. N. & Virshup, G. F. Superconducting oxide multilayers and superlattices: Physics, chemistry and nanoengineering. *Physica C* **235–240**, 178–181 (1994).
- Bozovic, I. *et al.* No mixing of superconductivity and antiferromagnetism in a high temperature superconductor. *Nature* **422**, 873–875 (2003).
- Gozar, A., Logvenov, G., Butko, V. B. & Bozovic, I. Surface structure analysis of atomically smooth BaBiO₃ films. *Phys. Rev. B* **75**, R201402 (2007).
- Sato, H., Tsukada, A., Naito, M. & Matsuda, A. La_{2-x}Sr_xCuO₄ epitaxial films ($x = 0$ to 2): Structure, strain, and superconductivity. *Phys. Rev. B* **61**, 12447–12456 (2000).
- Loquet, J.-P. *et al.* Doubling the critical temperature of La_{1.9}Sr_{0.1}CuO₄ using epitaxial strain. *Nature* **394**, 453–456 (1998).

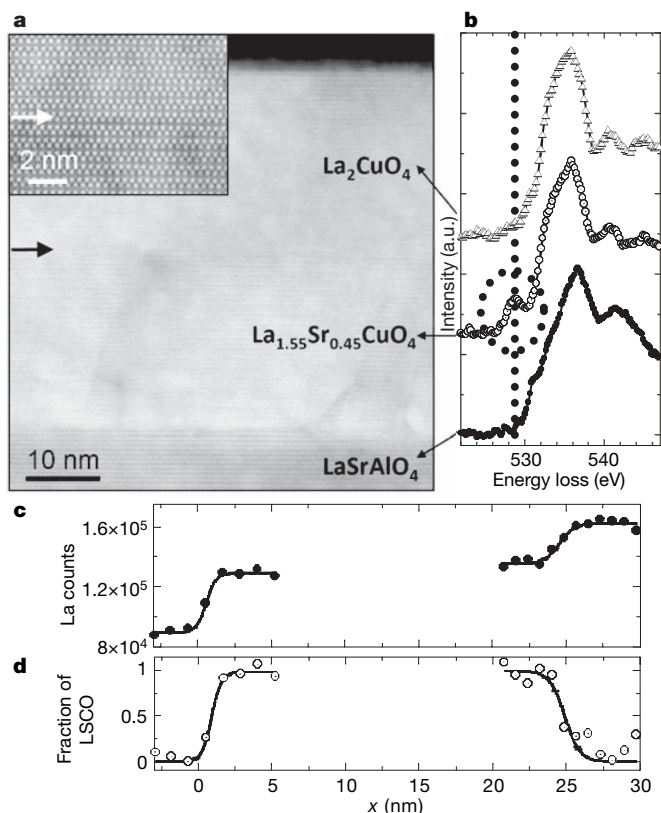


Figure 4 | Analysis of an M–I bilayer by scanning transmission electron microscopy and electron energy-loss spectroscopy. **a**, Annular dark field image of the structure (main panel). Inset, a magnified image of the M–I interface (arrowed). **b**, Oxygen-K (O-K) electron energy-loss spectra of the three oxides in the structure, showing clear changes in the fine structure of the O-K edge. For LSCO, an O-K edge pre-peak (circled) evolves for $x > 0$ and scales with the doping level^{26,27}. a.u., arbitrary units. **c**, The integrated La intensity across the bilayer. As expected, the La profile shows an increase in the La concentration from the substrate to the M layer, and again from the M to the I layer. **d**, Results of a principal-components analysis of the two interfaces. The fraction of one of two components, corresponding to the O-K edge in La_{1.55}Sr_{0.45}CuO₄, is shown.

16. Hebard, A. F. & Fiory, A. T. Evidence for the Kosterlitz-Thouless transition in thin superconducting aluminum films. *Phys. Rev. Lett.* **44**, 291–294 (1980).
17. Claassen, J. H., Reeves, M. E. & Soulen, R. J. Jr. A contactless method for measurement of the critical current density and critical temperature of superconducting rings. *Rev. Sci. Instrum.* **62**, 996–1004 (1991).
18. Clem, J. R. & Coffey, M. W. Vortex dynamics in a type-II superconducting film and complex linear-response functions. *Phys. Rev. B* **46**, 14662–14674 (1992).
19. Jensen, H. J. & Minnhagen, P. Two-dimensional vortex fluctuations in the nonlinear current-voltage characteristics for high-temperature superconductors. *Phys. Rev. Lett.* **66**, 1630–1633 (1991).
20. de Vries, J. W. C., Stollman, G. M. & Gijs, M. A. M. Analysis of the critical current density in high- T_c superconducting films. *Physica C* **157**, 406–414 (1989).
21. Smadici, S. *et al.* Hole delocalization in superconducting La_2CuO_4 - $\text{La}_{1.64}\text{Sr}_{0.36}\text{CuO}_4$ superlattices. Preprint at (<http://xxx.lanl.gov/abs/0805.3189>) (2008).
22. Ino, A. *et al.* Chemical potential shift in overdoped and underdoped $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **79**, 2101–2104 (1997).
23. Fujita, K., Noda, T., Kojima, K. M., Eisaki, H. & Uchida, S. Effect of disorder outside the CuO_2 planes on T_c of copper oxide superconductors. *Phys. Rev. Lett.* **95**, 097006 (2005).
24. Ginzburg, V. L. On interface superconductivity. *Phys. Lett.* **13**, 101–102 (1964).
25. Kivelson, S. A. Making high T_c higher: A theoretical proposal. *Physica B* **318**, 61–67 (2002).
26. Romberg, H., Alexander, M., Nücker, N., Adelman, P. & Fink, J. Electronic structure of the system $\text{La}_{2-x}\text{Sr}_x\text{CuO}_{4+\delta}$. *Phys. Rev. B* **42**, R8768–R8771 (1990).
27. Chen, C. T. *et al.* Electronic states in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_{4+\delta}$ probed by soft-x-ray absorption. *Phys. Rev. Lett.* **66**, 104–107 (1991).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The work at BNL was supported by US DOE. L.F.K. and D.A.M. acknowledge support under the ONR EMMA MURI and by the Cornell Center for Materials Research. L.F.K. acknowledges financial support by Applied Materials.

Author Contributions A.G. and G.L. contributed equally to this work. Film synthesis and characterization was by G.L. and I.B.; transport and time-of-flight ion scattering and recoil spectroscopy was by A.G.; lithography was by A.T.B.; electron microscopy was by L.F.K. and supervised by D.A.M.; and TEM sample preparation was by L.A.G. and L.F.K.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to I.B. (bozovic@bnl.gov).

LETTERS

Tibetan plateau river incision inhibited by glacial stabilization of the Tsangpo gorge

Oliver Korup¹ & David R. Montgomery²

A considerable amount of research has focused on how and when the Tibetan plateau formed in the wake of tectonic convergence between India and Asia¹. Although far less enquiry has addressed the controls on river incision into the plateau itself², widely accepted theory³ predicts that steep fluvial knick points (river reaches with very steep gradients) in the eastern Himalayan syntaxis at the southeastern plateau margin should erode rapidly⁴, driving a wave of incision back into the plateau. Preservation of the plateau edge thus presents something of a conundrum that may be resolved by invoking either differential rock uplift matching erosional decay^{5–7}, or other mechanisms for retarding bedrock river incision^{8,9} in this region where high stream power excludes the potential for aridity as a simple limit to dissection of the plateau¹⁰. Here we report morphologic evidence showing that Quaternary depression of the regional equilibrium line altitude, where long-term glacier mass gain equals mass loss, was sufficient to repeatedly form moraine dams on major rivers: such damming substantially impeded river incision into the southeastern edge of the Tibetan plateau through the coupled effects of upstream impoundment and interglacial aggradation. Such glacial stabilization of the resulting highly focused river incision centred on the Tsangpo gorge could further contribute to initiating and accentuating a locus of rapid exhumation, known as tectonic anaerism⁶.

Few studies have quantified the geomorphic consequences of glaciers that block mountain rivers beyond the perspective of hazard implications of catastrophic dam failure and outburst flows¹¹. Hence little is known about the potential topographic implications of glacial dams at a regional scale¹². Here we focus on the Yarlung Tsangpo and its major tributaries, the Yigong and Parlung Tsangpo that drain the southeastern margin of the Tibetan plateau near Namche Barwa (Fig. 1). In this area, where the Yarlung Tsangpo cuts through one of the deepest gorges on Earth, several independent methods—thermochronometric⁶, stream-power^{4,13}, and sediment provenance analyses¹⁴—reveal some of the highest erosion rates in the Himalaya, capable of exhuming young migmatitic gneisses since ~3.5 Myr ago⁶.

Montgomery *et al.*¹⁵ recognized evidence for repeated Holocene glacial damming of the Yarlung Tsangpo immediately above its gorge through the Himalaya, and similar blockages also occur on major tributaries upstream of moraine dams from deeply incised bedrock gorges. In this study, we mapped 260 moraine dams (Fig. 1), including numerous large laterofrontal moraines and associated debris fans, formed by glaciers flowing from tributary basins that blocked the Yarlung, Yigong and Parlung Tsangpo.

Topographic data on former glacier surfaces are not available for the study area, so we used the toe-to-summit-altitude method

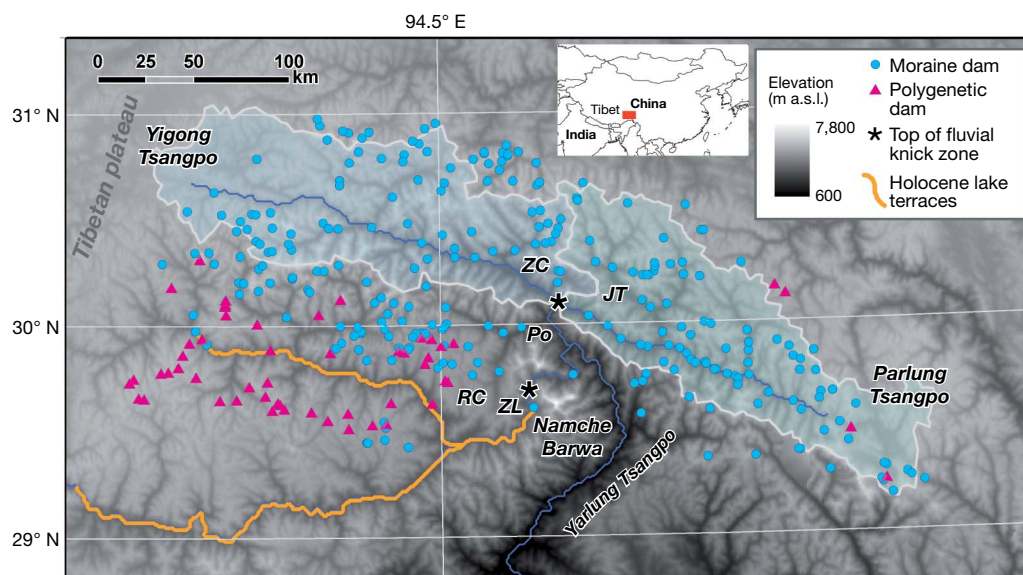


Figure 1 | Locations of 303 natural dams in the eastern Nyainqentanglha mountains and the Namche Barwa syntaxis, southeastern Tibetan plateau margin. JT, Jepu Tsangpo; Po, Po Tsangpo; RC, Rong Chu; ZC, Zhamu Creek landslide; ZL, Zelongnong glacier. Inset, location of study area.

Orange lines (see key) indicate estimated upstream extent of Holocene moraine-dammed lakes on the Yarlung Tsangpo¹⁵. Polygenetic dams composed mainly of glacial and landslide debris occur in fully deglaciated headwaters west of Rong Chu. a.s.l., above sea level.

¹Swiss Federal Research Institutes WSL/SLF, Flüelistr. 11, CH-7260 Davos, Switzerland. ²Quaternary Research Center, Department of Earth and Space Sciences, University of Washington, Seattle, Washington 98195-1310, USA.

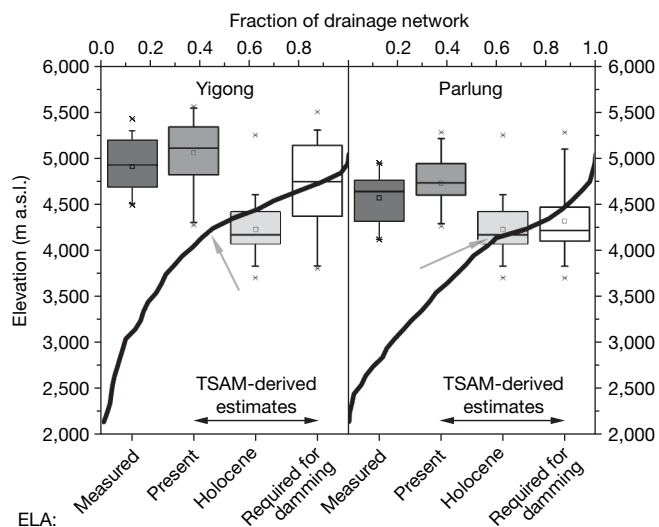


Figure 2 | Box-and-whisker plots of equilibrium line altitudes (ELAs) for Yigong and Parlung Tsangpo. ELAs are derived from field data^{16–19} and LANDSAT satellite imagery ('Measured'), and TSAM ('Present', 'Holocene', 'Required for damming'). TSAM systematically overestimates measured ELAs by ~200 m (Supplementary Information). TSAM-derived Holocene ELAs were inferred from locations of major moraine dams; also shown are ELAs required for tributary glaciers to dam trunk rivers where suspected moraine dams are poorly preserved. Solid line shows cumulative elevation distribution (hypsoetry) of each drainage network with contributing catchment areas >10 km², that is, exclusive of low-order channels and hillslopes. Arrows indicate glacial-fluvial transition. Box gives lower and upper quartiles and median; whiskers show 5th and 95th percentiles; small open square is mean, and crosses are 1st and 99th percentiles.

(TSAM) to calculate the likely range of depressions of the equilibrium line altitude, ΔELA , required to explain observed moraine dams and to dam major rivers. Advancing modern glaciers to the positions of 38 large laterofrontal moraines that blocked or constricted the Yigong, Parlung and Yarlung Tsangpo requires a mean ΔELA of 420 ± 170 m ($\pm 1\sigma$; Fig. 2). Sediments deposited in moraine-dammed lakes on the Yarlung Tsangpo¹⁵ document several episodes of substantial glacial advance throughout the Holocene, although the $\Delta ELAs$ associated with dated Little Ice Age (LIA) and Late Holocene (~1–2 kyr ago) moraines are smaller, that is, 10–60 m and 110–160 m, respectively¹⁶. From a sample of 89 basins with less direct evidence for glacial damming throughout the region, we computed that, on average, $\Delta ELA > 360$ m is required for glaciers from tributaries to dam trunk rivers. These estimates are remarkably consistent, despite the high local and regional variance in ELA due to local variations in the effects of topography, climate and supraglacial sediment flux.

The fraction of the drainage network length situated above a given moraine dam, and thus affected by glacial damming through direct occupancy and indirect impoundment or aggradation, increases with ΔELA . This fraction increases from a regional mean of 13% for the present ELA, to Holocene glacial advances that impounded between 30% and 55% of the length of the Yigong and Parlung Tsangpo, and commensurately higher for the greater ELA depression (which is currently debated, but may have amounted to ~1,000 m (refs 17, 18)) during earlier Pleistocene glacial advances. Hypsometric curves of river-channel elevations with contributing catchment areas >10 km² indicate that 40–60% of this drainage network may have been above our TSAM-derived Holocene ELA. Consequently, ELA depressions during both the Holocene and Pleistocene glacial maxima were more than adequate to block major rivers. The distinct kink in the drainage network hypsoetry at ~4,200 m coincides well

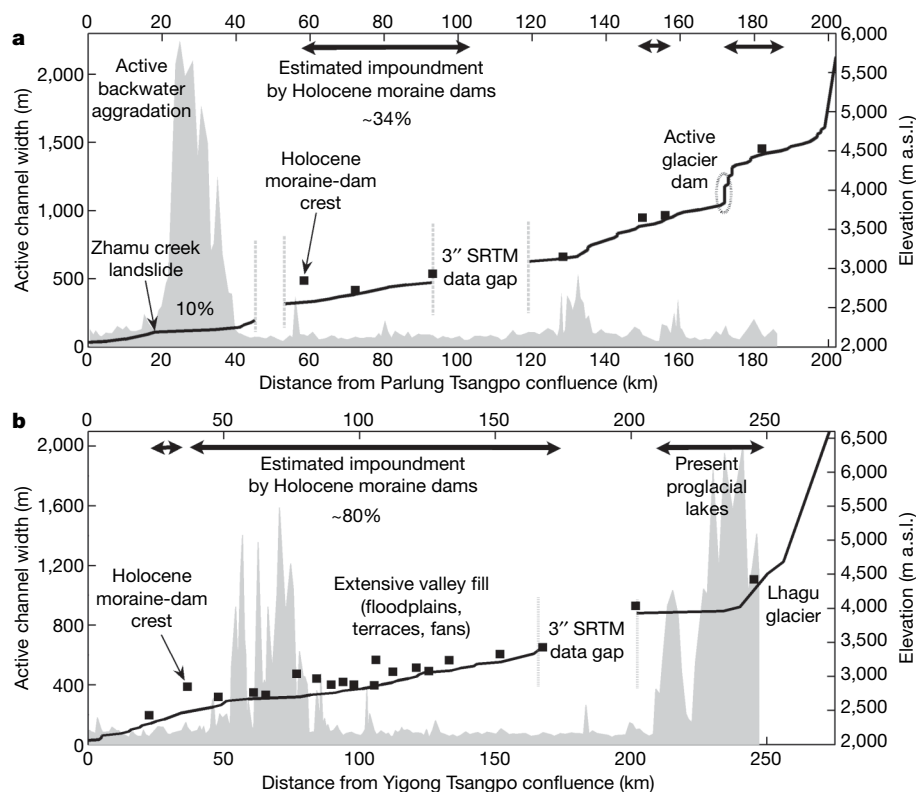


Figure 3 | Long profiles (lines) and active channel width (grey shaded areas) of Yigong Tsangpo and Parlung Tsangpo. a, Yigong Tsangpo; b, Parlung Tsangpo. Percentage of trunk-river length impounded by preserved Holocene moraine dams (extent shown by horizontal arrows) is estimated from horizontally projecting effective dam-crest elevations above

the present valley floor onto the river-long profile. Large historic landslides at Zhamu creek¹⁹ forced aggradation over 10% of trunk-river length. Large proglacial lakes inhibit bedrock incision in headwaters. Grey vertical dashed lines bracket SRTM data gaps.

with our reconstructions of the Holocene ELA, suggesting that either all the region's major river profiles reflect the influence of a sustained base level at elevations in the range of the Quaternary ELAs, or that headward knick point migration in these bedrock rivers has either stalled or not yet progressed any further into the Tibetan plateau.

Several major knick points in the long profiles of the Yigong and Parlung Tsangpo separate steep bedrock gorges from broad alluviated reaches with densely vegetated valley fills composed of floodplains, terraces and fans or wide, braided channels (Fig. 3). Although the most prominent of the active depositional 'valley trains' along the Yigong Tsangpo occurs upstream of an enormous modern landslide dam (Zhamu creek¹⁹), horizontal projections of the crests of major Holocene moraine dams onto the river-long profiles indicate that ~34% and ~80% of the Yigong and Parlung Tsangpo, respectively, were inundated or aggraded upstream of the dams. In many locations, the valley trains change to narrower single-thread channels where diverted around the toes of moraine dams, showing that the moraine dams exert a primary control on active channel width and thereby reduce specific stream power. Upstream of glacial dams, river incision into bedrock is dramatically impeded by impoundment during glacial occupancy. Moreover, glacial 'retreat' fosters blockage-induced backwater aggradation and infill of meltwater lakes with sediment. Even partial blockage by coarse moraine debris can control local base level through channel narrowing and armouring, backwater formation, and upstream alluviation. Hence, bedrock river incision below moraine dams contributes to further steepening river profiles; in contrast, glacier-controlled damming forces pervasive sediment trapping on valley floors upstream, thereby inhibiting bedrock river incision in upstream areas from somewhat below the ELA during interglacial periods up to about the ELA during glacial advances. Coupled numerical simulations of glacial and fluvial valley profile evolution show such inflections associated with the transition from glacial to fluvial bedrock incision²⁰.

The mapped pattern of moraine dams indicates that Holocene advances may have impeded bedrock incision over up to 80% of the trunk channel network. This appears realistic, given that the regional glacier extent was an estimated 30% larger during the LIA, and associated with blockage of major rivers, such as the Jipu Tsangpo²¹ (Fig. 1). Lake and backwater terraces extending over 560 km upstream associated with at least two Holocene glacier damming episodes of the Yarlung Tsangpo by the Zelongnong glacier, on the western flank of Namche Barwa, strikingly attest to widespread fluvial transport limitation forced by glacier advances¹⁵ (Fig. 1). This inundation affected 15–22% of the length of the Yarlung Tsangpo directly upstream of the moraine dams, and consequently contributed to retarding upstream knick point migration in this part of the plateau margin.

On all three rivers, channel-damming moraines extend downstream to the head of pronounced knick zones with inferred high stream power (Fig. 4). Moraine-dammed lakes not only inhibit bedrock incision during their lifetime, but also promote upstream aggradation and burial of bedrock valley floors under glaciofluvial sediment long after the dams are breached and the glaciers melt off. Hence, trunk river incision may be greatly reduced above these moraine dams, leaving glaciers alone responsible for excavating the valleys. Glaciers achieve most of their erosion near and above the ELA^{20,22}, and glacial over-excavation of these valleys will promote interglacial aggradation, thereby limiting the potential for further bedrock incision during interglacial periods. The 'fluvial' valleys downstream, however, will keep incising more rapidly, as long as episodic blockage or armouring by landslides^{19,23} is not as widespread, persistent or effective as that by glacial dams. Unlike moraine dams that repeatedly block the same location, landslide dams recur at variable locations along a river profile⁸. Hence, we propose that the juxtaposition of repeated glacier-controlled damming upstream and fluvial bedrock incision downstream contributes to stabilizing the position of the plateau edge at Namche Barwa by retarding the propensity for bedrock knick points to migrate upstream. This generates

a positive feedback, through which the rise of the Tibetan plateau to above the ELA would help preserve high topography, despite the innate tendency for river incision to sweep inland from steep plateau margins.

From a long-term perspective, we also find that when projecting published mineral cooling ages from the Namche Barwa region¹³ onto the closest point along the river profiles, the zone of youngest

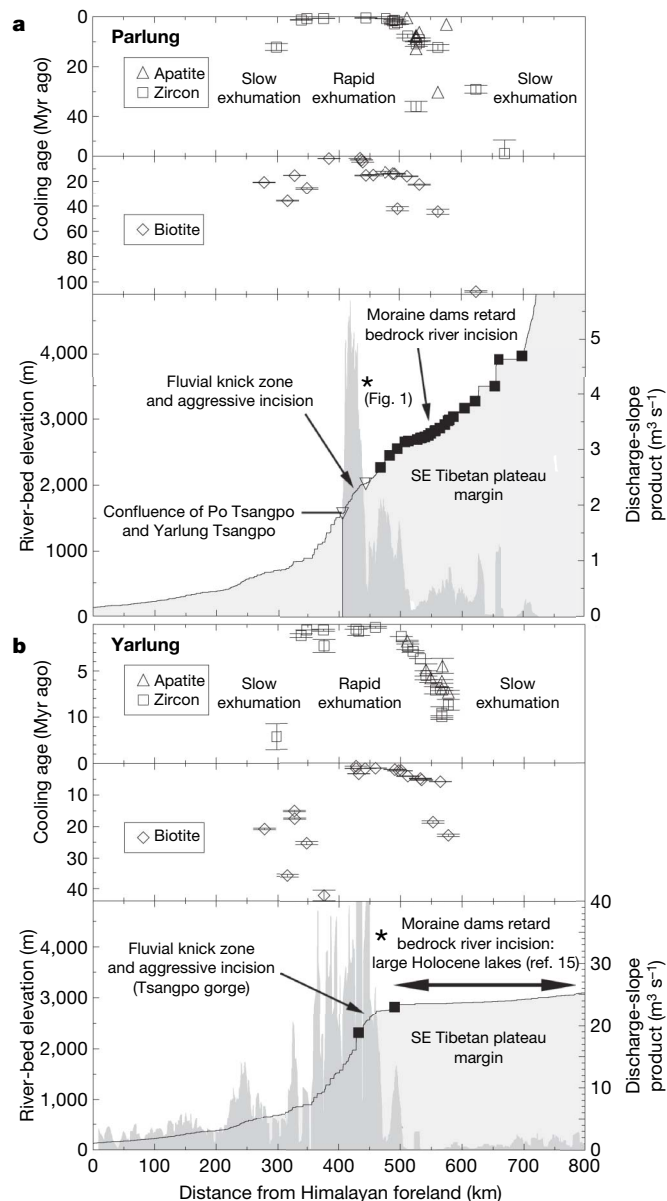


Figure 4 | Cooling ages (open symbols) and longitudinal profiles (dark lines) of the Parlung and Yarlung Tsangpo across the southeastern Tibetan plateau margin. a, Parlung Tsangpo; b, Yarlung Tsangpo. Profiles (lower panels) are measured from where the river leaves the Himalaya and enters the lowlands of northeastern India. They contain mapped locations of moraine dams (filled black squares), and local values of stream-power index (dark grey shading). This index is determined as the product of discharge and channel slope. Discharge is calculated using mean annual flow estimated as the area-weighted summation of local, TRMM-derived precipitation totals using data of ref. 27 and methods of ref. 4; channel slope is determined from a 10-km running average of local river elevation. Triangles on profile of the Parlung indicate confluences where the Yigong and Parlung merge to form the Yarlung Tsangpo (upstream), and where the Po Tsangpo enters the Yarlung Tsangpo (downstream). Asterisks show top of fluvial knick zones (see Fig. 1). Upper panels show projected locations of apatite, zircon and biotite cooling ages¹³ (error bars, ± 1 s.d.).

cooling ages (<1 Myr ago), and thus highest exhumation rates, extends downstream from the zone where 'string-of-pearls' clusters of moraine dams trap sediment and impede trunk-river incision into bedrock. On neither the Yarlung nor the Parlung Tsangpo is there evidence for a major offset in the dome-shaped pattern for different mineral cooling ages, as would be expected from substantial incision-driven headward migration of the Tsangpo knick points. However, new low-temperature chronology data not only confirm that the Namche Barwa antiform is <4 Myr old, but point to its gradual northeastward growth²⁴.

The proximity of repeated glacier blockage to rapid downstream fluvial incision that spatially focuses high erosive power means that drainage-blocking advances of temperate monsoonal valley glaciers from tributary channels could impede fluvial bedrock incision in trunk river systems. Whether driven by changes in monsoon precipitation^{18,25}, temperature^{16,17} or supraglacial sediment flux²⁶, Quaternary glacier advances appear capable of substantially retarding river incision into southeast Tibet in the region around Namche Barwa¹⁵. Although this is not the only viable mechanism for maintaining a plateau margin (as aridity¹⁰ and local fault offset⁷ provide viable complementary alternatives that are clearly important in some areas), our hypothesis frames a previously unrecognized and potentially important interaction between climate, tectonics and erosion in governing landscape evolution. In any case, glacial damming at the onset of mountain glaciation in the Quaternary may have produced a positive feedback that helped preserve a distinct plateau edge in close proximity to the extraordinary spatially focused exhumation of the Tsangpo gorge.

METHODS SUMMARY

Moraine dam mapping and ELA reconstruction. Moraine dams were identified and mapped from LANDSAT ETM+ satellite imagery. In the absence of data on former glacier surface topography, we used TSAM to reconstruct steady-state ELAs necessary for glacial advances preserved in conspicuous moraine ridges and loops. TSAM uses the arithmetic mean of the elevation difference between the highest summit peak and the lowest terminal moraine position for each catchment, determined from Space Shuttle Topographic Radar Mission (SRTM3) data at ~ 80 -m resolution.

TSAM results were calibrated with field-measured ELAs^{16–19} and estimates from LANDSAT (ETM+) imagery. To account for potential undersampling of moraine dams, we also used TSAM to estimate the Δ ELA required to advance glaciers to tributary junctions, and thereby block mainstem rivers.

Estimating fluvial erosion potential. We extracted river-long profiles and active channel widths from the SRTM and Landsat ETM+ imagery (band 8 at 15-m resolution), respectively. We mapped impounded or aggraded reaches upstream of present and former terminal moraines along the Yigong and Parlung Tsangpo. We calculated the cumulative elevation distribution (hypsometry) of this drainage network, that is, excluding all low-order channels and hillslopes in order to estimate the fraction of trunk drainage network located above the ELA for a given Δ ELA.

Discharge estimates were based on 4-yr average rainfall rates derived from Tropical Rainfall Measuring Mission (TRMM) data²⁷ and summed over the upstream drainage area for each point down a river profile. Channel slopes were computed as a running average over a 10-km-length reach centred on each point down the river profile using SRTM30 data.

We characterized inferred long-term (10^5 – 10^6 yr) erosion by projecting locations of published mineral (that is apatite, zircon (U-Th)/He, and biotite $^{40}\text{Ar}/^{39}\text{Ar}$) cooling ages¹³ onto the longitudinal profiles of the Parlung and Yarlung Tsangpo.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 January; accepted 30 July 2008.

1. Tapponnier, P. *et al.* Oblique stepwise rise and growth of the Tibet Plateau. *Science* **294**, 1671–1677 (2001).

2. Harkins, N., Kirby, E., Heimsath, A., Robinson, R. & Rieser, U. Transient fluvial incision in the headwaters of the Yellow River, northeastern Tibet, China. *J. Geophys. Res.* **112**, F03804, doi:10.1029/2006FJ000570 (2007).
3. Whipple, K. X. Bedrock rivers and the geomorphology of active orogens. *Annu. Rev. Earth Planet. Sci.* **32**, 151–185 (2004).
4. Finlayson, D. P., Montgomery, D. R. & Hallet, B. Spatial coincidence of rapid inferred erosion with young metamorphic massifs in the Himalayas. *Geology* **30**, 219–222 (2002).
5. Bendick, R. & Bilham, R. How perfect is the Himalayan arc? *Geology* **29**, 791–794 (2001).
6. Zeitler, P. K. *et al.* Erosion, Himalayan geodynamics, and the geomorphology of metamorphism. *GSA Today* **11**, 4–9 (2001).
7. Lavé, J. & Avouac, J. P. Fluvial incision and tectonic uplift across the Himalayas of central Nepal. *J. Geophys. Res.* **106**, 26561–26591 (2001).
8. Korup, O., Strom, A. L. & Weidinger, J. T. Fluvial response to large rock-slope failures: Examples from the Himalayas, the Tien Shan, and the Southern Alps in New Zealand. *Geomorphology* **78**, 3–21 (2006).
9. Ouimet, W. B., Whipple, K. X., Royden, L. R., Sun, Z. & Chen, Z. The influence of large landslides on river incision in a transient landscape: Eastern margin of the Tibetan Plateau (Sichuan, China). *Geol. Soc. Am. Bull.* **119**, 1462–1476 (2007).
10. Sobel, E. R., Hilley, G. E. & Strecker, M. R. Formation of internally drained contractional basins by aridity-limited bedrock incision. *J. Geophys. Res.* **108** (B7), 2344, doi:10.1029/2002JB001883 (2003).
11. Clague, J. J. & Evans, S. G. A review of catastrophic drainage of moraine-dammed lakes in British Columbia. *Quat. Sci. Rev.* **19**, 1763–1783 (2000).
12. Korup, O. & Tweed, F. S. Ice, moraine, and landslide dams in mountainous terrain. *Quat. Sci. Rev.* **26**, 3406–3422 (2007).
13. Finnegan, N. J. *et al.* Coupling of rock uplift and river incision in the Namche Barwa-Gyala Peri massif, Tibet. *Geol. Soc. Am. Bull.* **120**, 144–152 (2008).
14. Garzanti, E. *et al.* Sand petrology and focused erosion in collision orogens: The Brahmaputra case. *Earth Planet. Sci. Lett.* **220**, 157–174 (2004).
15. Montgomery, D. R. *et al.* Evidence for Holocene megafloods down the Tsangpo River gorge, southeastern Tibet. *Quat. Res.* **62**, 201–207 (2004).
16. Yang, B., Bräuning, A., Dong, Z., Zhang, Z. & Jiao, K. Late Holocene monsoonal temperate glacier fluctuations on the Tibetan Plateau. *Glob. Planet. Change* **60**, 126–140 (2008).
17. Shi, Y. Characteristics of late Quaternary monsoonal glaciation on the Tibetan Plateau and in East Asia. *Quat. Int.* **97–98**, 79–91 (2002).
18. Owen, L. A. & Benn, D. I. Equilibrium-line altitudes of the Last Glacial Maximum for the Himalaya and Tibet: An assessment and evaluation of results. *Quat. Int.* **138–139**, 55–78 (2005).
19. Shang, Y. *et al.* A super-large landslide in Tibet in 2000: Background, occurrence, disaster, and origin. *Geomorphology* **54**, 225–243 (2003).
20. Anderson, R. S., Molnar, P. & Kessler, M. A. Features of glacial valley profiles simply explained. *J. Geophys. Res.* **111**, F01004, doi:10.1029/2005FJ000344 (2006).
21. Su, Z. & Shi, Y. Response of monsoonal temperate glaciers to global warming since the Little Ice Age. *Quat. Int.* **97–98**, 123–131 (2002).
22. Mitchell, S. G. & Montgomery, D. R. Influence of a glacial buzzsaw on the height and morphology of the central Washington Cascade Range, USA. *Quat. Res.* **65**, 96–107 (2006).
23. Korup, O. Rock-slope failure and the river long profile. *Geology* **34**, 45–48 (2006).
24. Seward, D. & Burg, J. P. Growth of the Namche Barwa Syntaxis and associated evolution of the Tsangpo Gorge: Constraints from structural and thermochronological data. *Tectonophysics* **451**, 282–289 (2008).
25. Wang, Y. *et al.* The Holocene Asian monsoon: links to solar changes and North Atlantic climate. *Science* **308**, 854–857 (2005).
26. Hewitt, K. Catastrophic landslide deposits in the Karakoram Himalaya. *Science* **242**, 64–67 (1988).
27. Anders, A. M. *et al.* In *Tectonics, Climate and Landscape Evolution* (eds Willett, S. D., Hovius, N., Brandon, M. T. & Fisher, D. M.) 39–54 (Geological Society of America Special Paper 398, GSA, Washington DC, 2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements O.K. was partly supported by EU-FP6 contract 081412 IRASMOS. D.R.M. acknowledges support from the Continental Dynamics Program of the US National Science Foundation (EAR-0003561). We thank H. Greenberg for assistance in figure drafting. Reviews by D. Burbank and L. Owen helped improve an earlier manuscript.

Author Contributions Both authors contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to O.K. (korup@slf.ch).

METHODS

Moraine dam mapping and ELA reconstruction. Moraine dams were identified from LANDSAT ETM+ satellite images on the basis of a suite of morphologic characteristics, including the presence of glaciers in headwaters and concomitant occurrence of impounded lakes. The single, sharp and mostly vegetated crests indicate little modification by post-depositional surface processes²⁸, while ruling out deposition as landslide debris during catastrophic emplacement because of a lack of characteristic morphology²⁹. There are also numerous remnants of terminal moraines that appear to have been truncated by fluvial erosion along narrow bedrock gorges, particularly on the Yigong and the Yarlung Tsangpo. Moreover, the proximity of modern glaciers to mainstem rivers suggests that additional glacial dams probably occurred in numerous locations where scant or no direct morphologic evidence speaks to their former presence. We have not included such tentatively identified potential glacial dams in our regional map (Fig. 1). Therefore, the number of moraine dams mapped in Fig. 1 provides a minimum estimate of the overall effects of glacial damming.

In the absence of data on former glacier surface topography, we used TSAM to reconstruct ELAs necessary to produce the advances preserved in conspicuous moraine ridges and loops. TSAM uses the arithmetic mean of the elevation difference between the highest summit peak and the lowest terminal moraine position for each catchment (for a thorough discussion of this and alternative methods of reconstructing former ELAs, see ref. 30). It is ideal for reconstructing former ELAs at a regional scale from 3 arcsec Space Shuttle Topographic Radar Mission (SRTM3) digital elevation data. These data suffer from local data gaps in this region of high local relief and have been merged with SRTM30 data to provide a nominal grid-cell resolution of ~80 m.

In order to calibrate our TSAM results, we obtained the locations of present ELA from the literature^{16–19} and measurements from LANDSAT and LANDSAT ETM+ imagery at the end of the monsoon season. The annual ELAs on these images are closely approximated by clearly detectable, conspicuous colour differences between ice and firn cover. The monsoonal temperate glaciers of south-east Tibet are very sensitive to climatic change, as small temperature changes lead to substantial response in the glacier mass balance¹⁷. We compared multitemporal satellite images to confirm a negative mass balance for 20 of the larger glaciers between 1987 and 2001.

TSAM helps approximate the steady-state ELA for a given glacier. Given that the field-derived positions of the ELA were acquired during a period of negative mass balance, we can assume that these values somewhat, although probably slightly, overestimate the position of the steady-state ELA. Comparison of measured modern ELA positions versus TSAM-derived modern ELA positions reveals that TSAM systematically overestimates measured ELA by ~200 m, and is less reliable for determining ELA on debris-covered glaciers and those prone to surging in the study area¹⁷. Nonetheless, TSAM provides conservative (that is, minimum) estimates with respect to necessary conditions for river blocking, as surging will more readily impound rivers with climatic changes involving even less ELA depression.

We focused on glaciers issuing from mostly steep tributary basins with areas between 10 and 260 km² and a total relief of between 1.2 and 5.4 km. In order to account for potential erosional censoring of moraine dams, especially at sites

with limited geomorphic evidence of former blockage, we also used TSAM to estimate the minimum Δ ELA required to advance glacier snouts to tributary junctions, and thereby block mainstem rivers.

Estimating fluvial erosion potential. We extracted river-long profiles and active channel widths from the SRTM3 data and Landsat ETM+ images (monochromatic band 8 at 15-m resolution), respectively. We also mapped the length of the Yigong and Parlung Tsangpo either impounded or aggraded upstream of present and former terminal moraine positions, as well as the ice-covered region upstream of the tributary. Because of the available DEM resolution, we used an upstream area >10 km² to automatically delineate and measure the length of major river branches. This roughly corresponds to the length of fluvial channels, as the transition from debris-flow dominated colluvial channels to fluvially incised bedrock rivers typically occurs at drainage areas of 1–10 km² (ref. 31). While this transition arguably varies between individual catchments, there is documented evidence that debris flows are frequent geomorphic agents in many of the lower-order tributary basins³². We then calculated the cumulative elevation distribution (hypsoetry) of this drainage network, that is, excluding all low-order channels and hillslopes, for both the Yigong and Parlung Tsangpo in order to estimate the fraction of trunk drainage network located above the ELA for a given Δ ELA. We note that this value does not necessarily reflect the fraction of landscape that was covered by glaciers; we are here interested in reconstructing the potential geomorphic effects of glaciers on river systems rather than the total extent of Pleistocene glaciation.

Discharge estimates were based on 4-yr average rainfall rates derived from TRMM data²⁷ and summed over the upstream drainage area for each point down a river profile. We have chosen this approach to include the effects of the regional precipitation gradient across the Tsangpo-Brahmaputra catchment. Channel slopes were computed as a running average over a 10-km-length reach centred on each point down the river profile using SRTM30 data.

Finally, we characterized inferred long-term (10^5 – 10^6 yr) erosion by projecting the locations of published mineral (that is apatite, zircon (U-Th)/He, and biotite ⁴⁰Ar/³⁹Ar) cooling ages¹³ onto the nearest locations of the river-long profiles of the Parlung and Yarlung Tsangpo. We chose this approach in order to test whether the locations of the reconstructed glacier fluctuations and their effect on retarding fluvial bedrock incision would be detectable in the long-term development of the Tibetan plateau margin.

28. Lliboutry, L., Morales Arno, B., Pautre, A. & Schneider, B. Glaciological problems set by the control of dangerous lakes in Cordillera Blanca, Peru. I. Historic failures of morainic dams, their causes and prevention. *J. Glaciol.* **18**, 239–254 (1977).
29. Cruden, D. M. & Varnes, D. J. in *Landslides, Investigation and Mitigation* (eds Turner, A. K. & Schuster, R. L.) 36–75 (Transport Research Board, National Research Council, Washington DC, 1996).
30. Benn, D. I. & Lehmkuhl, F. Mass balance and equilibrium-line altitudes of glaciers in high-mountain environments. *Quat. Int.* **65–66**, 15–29 (2002).
31. Brummer, C. J. & Montgomery, D. R. Downstream coarsening in headwater channels. *Wat. Resour. Res.* **39**, 1294, doi:10.1029/2003WR001981 (2003).
32. Cheng, Z., Wu, J. & Geng, X. Debris flow dam formation in southeast Tibet. *J. Mount. Sci.* **2**, 155–163 (2005).

LETTERS

Central role of detachment faults in accretion of slow-spreading oceanic lithosphere

J. Escartín¹, D. K. Smith², J. Cann³, H. Schouten², C. H. Langmuir⁴ & S. Escrig⁴

The formation of oceanic detachment faults is well established from inactive, corrugated fault planes exposed on sea floor formed along ridges spreading at less than 80 km Myr⁻¹ (refs 1–4). These faults can accommodate extension for up to 1–3 Myr (ref. 5), and are associated with one of the two contrasting modes of accretion operating along the northern Mid-Atlantic Ridge. The first mode is asymmetrical accretion involving an active detachment fault⁶ along one ridge flank. The second mode is the well-known symmetrical accretion, dominated by magmatic processes with subsidiary high-angle faulting and the formation of abyssal hills on both flanks. Here we present an examination of ~2,500 km of the Mid-Atlantic Ridge between 12.5 and 35° N, which reveals asymmetrical accretion along almost half of the ridge. Hydrothermal activity identified so far in the study region is closely associated with asymmetrical accretion, which also shows high levels of near-continuous hydroacoustically and teleseismically recorded seismicity. Increased seismicity is probably generated along detachment faults that accommodate a sizeable proportion of the total plate separation. In contrast, symmetrical segments have lower levels of seismicity, which occurs primarily at segment ends. Basalts erupted along asymmetrical segments have compositions that are consistent with crystallization at higher pressures than basalts from symmetrical segments, and with lower extents of partial melting of the mantle. Both seismic evidence and geochemical evidence indicate that the axial lithosphere is thicker and colder at asymmetrical sections of the ridge, either because associated hydrothermal circulation efficiently penetrates to greater depths or because the rising mantle is cooler. We suggest that much of the variability in sea-floor morphology, seismicity and basalt chemistry found along slow-spreading ridges can be thus attributed to the frequent involvement of detachment faults in oceanic lithospheric accretion.

The large fields of detachment surfaces recently identified in oceanic crust formed along the slow-spreading Mid-Atlantic Ridge (MAR) and ultra-slow spreading South-West Indian Ridge^{3,6} demonstrate that these structures are involved in the accretion of a larger portion of the oceanic lithosphere than previously inferred from sea-floor corrugated planes alone⁷. The resulting sea-floor morphology and lithospheric structure on the conjugate flanks of the ridge axis are strongly asymmetrical³ and differ from the more regular and roughly symmetrical axis-parallel abyssal hill fabric believed to characterize both flanks of normal slow-spreading sea floor. The abyssal hill morphology is caused by ridge-parallel, high-angle faulting of volcanic sea floor⁸ (Fig. 1a–c). In contrast, detachment-related terrain is caused by long-lived steep, normal faults initiated beneath the rift valley floor that rotate to low angles as their footwalls are exposed^{7,8}. Distinctive narrow ridges with steep outward-facing slopes that are often curved in plan view develop near to exposed detachments at the sea floor, and

form the boundaries to deep swales⁷ (Fig. 1d, e), producing blocky and chaotic terrain^{7,9}. The asymmetric nature of accretion in the presence of detachments is also observed in the overall lithospheric structure, composition and geophysical character wherever data are available^{3,4,10}. Detachments at the MAR do not form the broad ridges without striations that are found along the ultra-slow spreading South-West Indian Ridge³.

Multibeam bathymetry data are available for most of the northern MAR and its flanks between the Marathon and the Oceanographer fracture zones (12° 40' N to 35° 15' N, Fig. 1a), a distance of ~2,500 km. We use the systematic morphological differences between abyssal hill terrain (Fig. 1b, c) and detachment-related terrain⁷ (Fig. 1d, e) to re-interpret the existing bathymetry on both flanks of this stretch of ridge. We quantify the importance of detachment faulting in lithospheric accretion, and investigate relationships between the mode of accretion and seismic character of the spreading axis and geochemistry of erupted basalts (Figs 2 and 3). Our analysis shows that symmetrical segments with abyssal hill terrain flanking both sides and no detachment faulting occupy more than 30% of the axis. Asymmetrical accretion where detachment faulting occurs along one flank of the axis makes up ~50% of the ridge axis. The remaining ~20% of the axis is unclassified owing to limited bathymetric coverage or to unclear morphology at discontinuities and oblique ridge sections (see interpreted maps and Supplementary Table 1). Although there is no apparent correlation with axial depth (Fig. 2a), nor a continuous along-axis trend in the relative distribution of accretionary modes, there is a broad wavelength change. Detachments are dominant between the Marathon and Fifteen-Twenty fracture zones, with 70% of the ridge axis accreting asymmetrical, and they are practically absent between the Hayes and Oceanographer fracture zones, immediately south of the Azores hotspot, where less than 15% of the axis accretes asymmetrical (Fig. 2a).

There is an excellent correlation between mode of accretion and seismicity at the ridge axis. This section of the MAR was hydroacoustically monitored between January 1999 and September 2003¹¹. The hydroacoustic catalogue is complementary to the NEIC teleseismic catalogue from 1973 to 2007 (see Methods), as it records smaller-magnitude events (magnitude of completeness of 3 and 5, respectively¹²), over a shorter period of time (<5 rather than >30 yr). The two seismic catalogues show that detachment-dominated, asymmetrical ridge sections host >50% more hydroacoustic and teleseismic events than symmetrical segments (Fig. 2b, c). The concentration of seismicity at segments shown to have active detachment faults (Fig. 1d, e), such as the Logachev massif south of the Fifteen-Twenty fracture zone and the TAG detachment fault near 26° N (refs 6,7,13), is thus a general pattern. Active detachments also control the zones of sustained seismicity, which lack shock-aftershock sequences that were previously identified along the northern MAR¹⁴. Differences between

¹Marine Geosciences Group, CNRS Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris Cedex 05, France. ²Geology and Geophysics Department, Woods Hole Oceanographical Institute, Woods Hole, Massachusetts 02543, USA. ³School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK. ⁴Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.

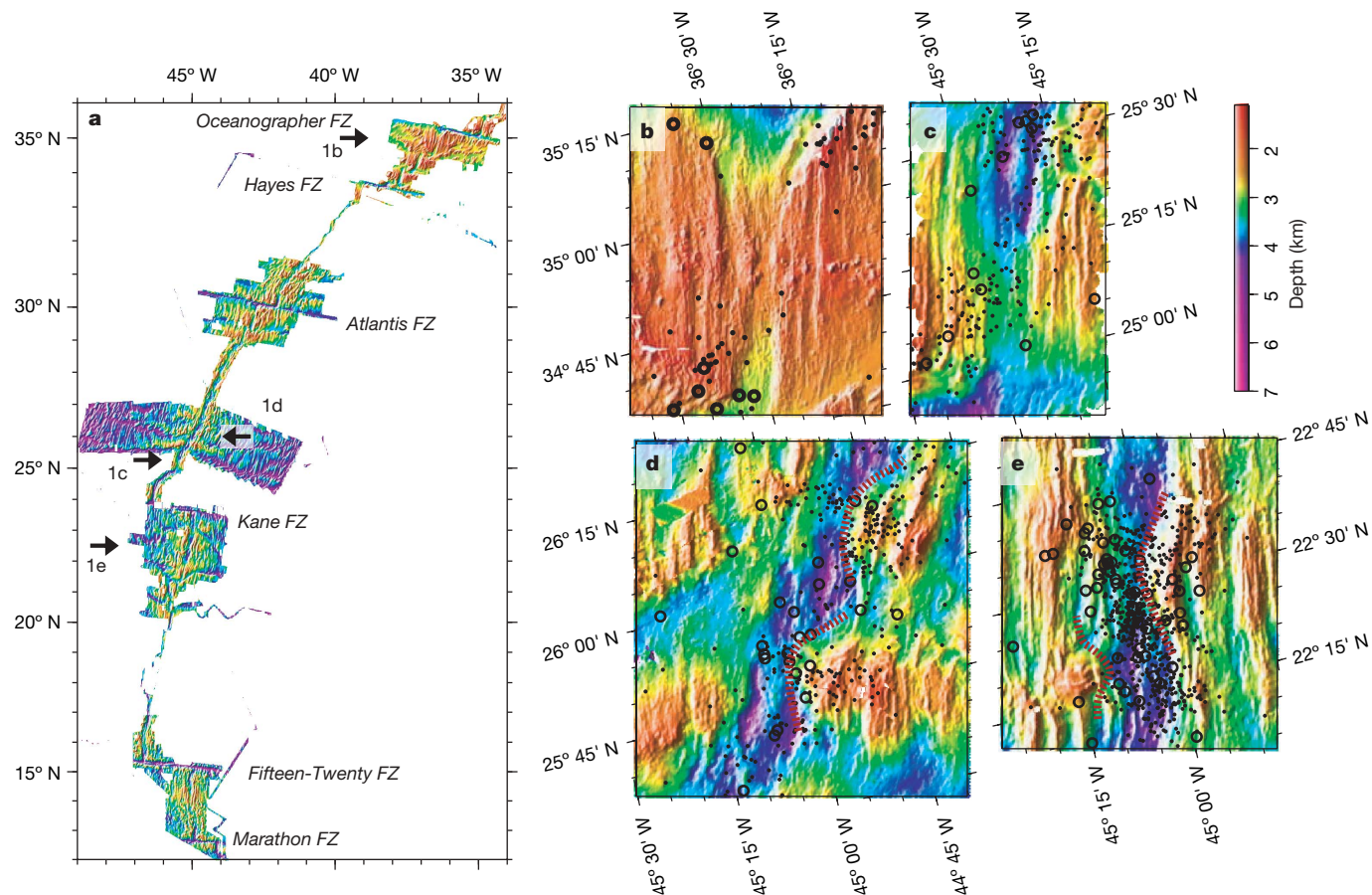


Figure 1 | Bathymetry of the study area and examples of symmetrical and asymmetrical segments with associated seismicity. **a**, Available multibeam bathymetric coverage data between the Marathon and Oceanographer fracture zones^{6,9,28}. **b**, **c**, Examples of linear ridge segments flanked by ridge-parallel abyssal hills resulting from high-angle normal faulting corresponding to symmetrical accretion at their centre. Seismicity

concentrates at segment ends, whereas the centre shows no or very few events. **d**, **e**, Examples of asymmetrical accreting ridge sections. The axis is flanked by active detachment faults, which are associated with elevated seismicity rates. Dashed red lines (**d** and **e**): termination of detachment towards the ridge axis; dots: hydroacoustic events¹¹; open circles: teleseismic events.

the hydrophone and teleseismic patterns (Fig. 2b, c) probably arise from temporal variability in seismicity, and from a secondary control on seismicity by other factors in addition to detachment faulting. For example, shock–aftershock sequences generated along steep normal faults in the flanks of the MAR may result in higher rates of hydrophone events along symmetrical segments, such as those near 24.25° N¹², which are not observed in the longer-term teleseismic data (Fig. 2b, c). Complex tectonic processes at segment ends and oblique zones can also significantly increase seismicity, even if segments show little or no seismicity at their centres (Fig. 1b, c)¹⁴.

Comparison of microseismicity results with regional seismic patterns allows us to constrain the nature of the zones of elevated and sustained seismicity. An 8-month microseismic experiment at the TAG hydrothermal field (26.1° N) shows deformation along the active detachment fault accommodated by continuous creep associated with steady hydroacoustic event rates (Fig. 1d). Microseismicity events at ridge sections with active detachments (TAG^{13,15}, 23° N¹⁶) are observed down to 7–8 km below sea floor, with the interpreted fault rooting directly below the neovolcanic zone¹³. In contrast, at the centre of symmetrical segments (35° N¹⁷ and Fig. 1b, 29° N¹⁸) events extend to maximum depths of 5–6 km, 1–3 km shallower than near detachments (see Supplementary Fig. 2).

There is a close association of hydrothermal activity and asymmetrical accretion. Ridge sections flanked by active detachments host seven out of the eight known active or recent hydrothermal vent fields in the area, and are overlain by six out of seven identified hydrothermal

plumes¹⁹ in the water column (Fig. 2a). There is also evidence for high-temperature fluid circulation along detachment faults^{20,21}, suggesting a link between hydrothermal activity and deformation associated with the faults. The data thus point to a link, albeit complex, between extension on detachments and persistent hydrothermal activity, consistent with a thicker lithosphere at asymmetrical than at symmetrical ridge sections.

The two modes of lithospheric accretion from our study area also show differences in the geochemistry of basalts. Trace element and to some extent major element data are subject to regional mantle heterogeneities²², in addition to local variability²³. We thus restrict our investigation to pairs of adjacent segments, one of which has asymmetrical and the other symmetrical accretion, and from both of which multiple chemical analyses are available. We note, however, that the trends described here are present when all data from the study area are considered, and despite significant overlap between basalt compositions from each of the two modes of accretion (see Supplementary Fig. 3). Basalts from symmetrical segments show less compositional variation, and generally reflect lower eruption temperatures, as expressed by the MgO contents of the lavas. Asymmetrical segments yield more primitive (higher-temperature) basaltic compositions, having undergone less crystal fractionation than basalts from symmetrical segments. There are also chemical offsets at constant MgO for major elements. Samples from asymmetrical segments have higher Na₂O and FeO and lower CaO than samples from symmetrical segments (Fig. 3a–c).

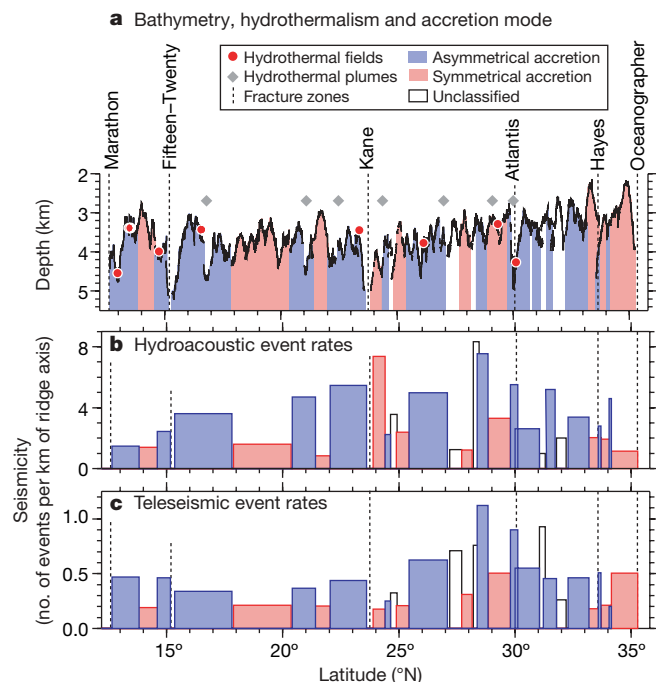


Figure 2 | Along-axis distribution of asymmetrical and symmetrical ridge sections and correlation with hydrothermal and seismic activity. **a**, Along-axis bathymetric profile showing the two modes of lithospheric accretion interpreted from sea-floor morphology and tectonic structure (see detailed interpreted maps in Supplementary Fig. 1). Segments associated with active detachment faulting host the large majority of known vent fields (red circles; see Methods) and identified hydrothermal plumes in the water column (grey diamonds)¹⁹. **b**, **c**, The average frequency of hydroacoustic (**b**; January 1999 to August 2003) and teleseismic events (**c**; January 1972 to December 2007) per kilometre of ridge axis is higher along asymmetrical than symmetrical segments.

Because there are only limited trace element data for adjacent symmetrical/asymmetrical ridge sections that can be used to compare chemical compositions, we can only compare the areas north and south of the Kane fracture zone (north of Kane, symmetrical, known as MARNOK; south of Kane, asymmetrical, known as MARK). For this pair of segments we compare the ratio of moderately incompatible elements, Zr/Y, because this ratio is not particularly influenced by mantle heterogeneity. The symmetrical MARNOK segment has lower and more homogeneous Zr/Y ratios than the asymmetrical MARK segment (Fig. 3d).

As shown by the liquid lines of descent due to cooling at different pressures (Fig. 3), the variations in FeO and CaO contents are consistent with lower-pressure fractionation for samples from symmetrical segments, and higher-pressure fractionation for samples from asymmetrical segments. The greater homogenization of lavas at symmetrical segments could be achieved within magma reservoirs resulting from greater melt supply to lower pressures than at asymmetrical ridge sections. In addition to the pressure of crystallization, other processes near detachment faults such as melt–rock reactions inferred from gabbros sampled along the Kane detachment²⁴ may also contribute to the systematic compositional difference between basalts erupted along the two types of segments.

High-pressure fractionation alone, however, does not fully account for the offset in Na₂O contents, nor does it produce significant differences in the Zr/Y ratio, which may result instead from changes in the extent of melting, or reflect mantle heterogeneities. Indeed, there are significant differences in incompatible element compositions of the two segments, with Ba contents, for example, varying by a factor of 2, and Ba/Y ratios in MARNOK of 0.4 ± 0.1 and in MARK of 0.23 ± 0.06 . Since the MARNOK segment has the higher

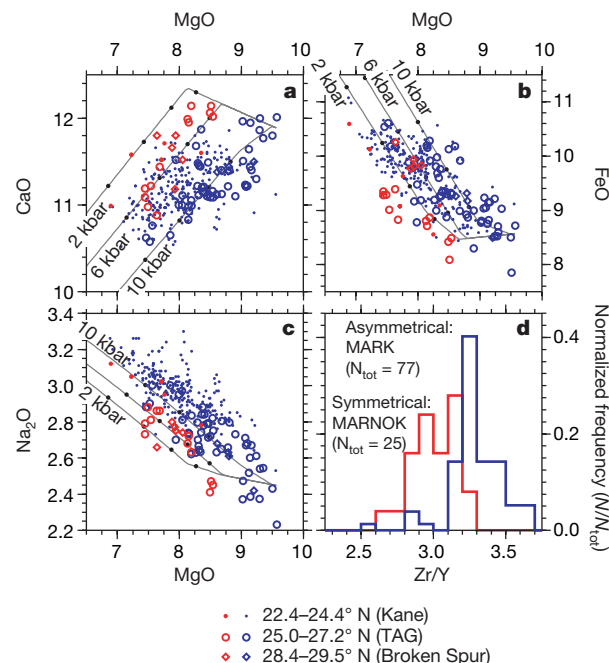


Figure 3 | Systematic differences in basalt chemistry from symmetrical and asymmetrical ridge sections. Blue symbols, symmetrical segments; red, asymmetrical segments. The systematic differences in MgO versus CaO (**a**), FeO (**b**) and Na₂O content (**c**) for three pairs of adjacent symmetrical and asymmetrical segments, and lines of liquid descent (black lines) showing the compositional evolution of basalts for different pressures of crystallization. Sufficient rare-earth element data are only available for the ridge sections adjacent to the Kane fracture zone, MARNOK to the north (symmetrical) and MARK to the south (asymmetrical). The Zr/Y ratio (**d**) is higher for the southern MARK segment.

Ba/Y ratio, it might be expected to have a higher Zr/Y ratio, Zr being less compatible than Y. However, Zr/Y is lower in MARNOK than in MARK (Fig. 3), the opposite of what would be predicted from a mantle heterogeneity effect. These data thus suggest melting differences between symmetrical and asymmetrical segments. Additional basalt analyses from other segment pairs are necessary to elucidate these observations further.

Accretion along slow-spreading ridges is thus dominated by two distinct modes of spreading and partitioning of deformation, which in turn are reflected in the thermal structure, magmatic system and hydrothermal circulation at the axis (Fig. 4). Accretion in the presence of detachments is highly asymmetrical, with ~50% of the plate separation accommodated along a single fault along one of the ridge flanks. Overall tectonic strain at symmetrical segments is probably much lower (<20%)^{25–27}, and distributed instead over multiple high-angle normal faults that are active for several tens of kilometres off-axis¹⁴. This difference in the amount of tectonic strain may be responsible for the high rates of seismicity at asymmetrical ridge sections relative to symmetrical ones. Increased hydrothermal cooling and lithospheric thickness in the presence of detachment faults is likely to influence the underlying magmatic system. A deepening of the top of the melting regime can lead to a reduction of the extent of melting of the mantle. Once magma separates and ascends towards the surface, it cools, reacts and crystallizes. There is thus a natural association between the thicker lithosphere of asymmetrical segments, the lower extents of melting, the higher degree of melt–rock reactions, and the greater pressures of crystallization at asymmetrical segments compared to symmetrical segments. One important remaining problem is that of the origin of the detachment faulting. Do detachment faults establish because of greater hydrothermal cooling and fault weakening that leads to the development of long-lived faults? Or are they forced to form by lower temperatures of the rising

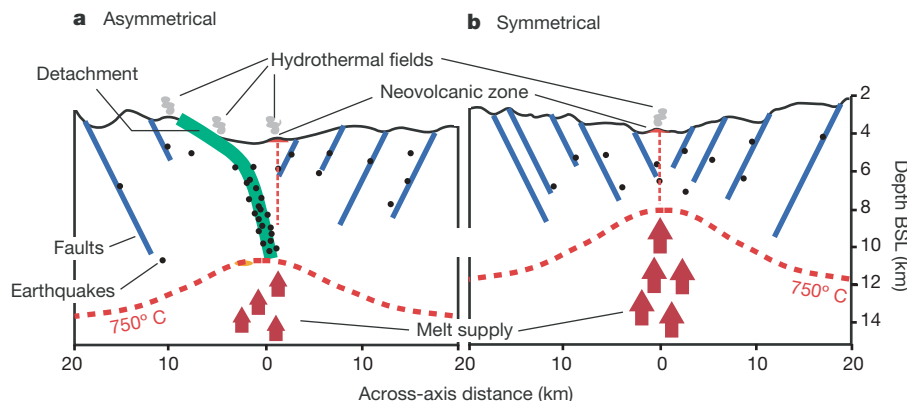


Figure 4 | Across-axis sections corresponding to symmetrical and asymmetrical accretion and associated processes. **a**, In asymmetrical ridge sections, detachments accommodate roughly half of the plate separation tectonically along a single fault. They also host hydrothermal fields that can contribute to the thick lithosphere at the axis (up to 8 km) inferred from microseismic data. This strain partitioning results in higher and more sustained levels of seismicity than symmetrical segments (**b**), which

accommodate a lower proportion of the plate spreading by faulting (<20%) over numerous faults on both flanks simultaneously, and in a thinner axial lithosphere (<6 km). The colder axial regime in asymmetrical segments can reduce the melting column, favour reduced melt supply and promote crystallization at deep levels relative to symmetrical segments. Alternatively, mantle anomalies may instead cause these along-axis variations in thermal regime and mode of accretion. BSL, below sea level.

mantle beneath the axis, and consequently decreased magma supply? There may be a positive feedback between these two effects, where increased cooling at the surface leads to decreased melt production at depth, and vice versa, causing small differences to be amplified and thus generate widely different styles of lithospheric accretion.

METHODS SUMMARY

Multibeam bathymetry data^{6,7,28} were used to identify the rideward termination of detachment terrain, and the extent of ridge-parallel abyssal hill terrain flanking the axis. Ridge sections were classified based on the presence of linear abyssal hill terrain on both flanks (symmetrical accretion), or of detachment terrain along one of the flanks (asymmetrical accretion). The remaining sections of the ridge axis were unclassified.

Seismicity data were obtained from the hydrophone catalogue¹¹ and from the NEIC teleseismic catalogue (<http://neic.usgs.gov/>). To remove intraplate seismicity and that associated with off-axis deformation, we restricted our analysis to events in oceanic lithosphere younger than 3 Myr, based on the digital sea-floor age map²⁹. Seismicity rates were calculated from the total number of hydroacoustic (Fig. 2b) and teleseismic events (Fig. 2c) recorded, and the length of individual ridge sections showing symmetric, asymmetric, or undetermined accretion.

Geochemical data for basalt glass were obtained from the PETDB database³⁰ and complemented with unpublished data (J. Cann and C. Langmuir). For the data analysis we selected those basalt analyses from rocks sampled along the axial volcanic ridge, the rift valley floor or the flanking walls. We then selected those analyses that could be corrected for interlaboratory bias³¹, discarding the remaining data. Liquid lines for different pressures of crystallization shown in Fig. 3 were calculated using an initial primitive basalt composition.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 March; accepted 6 August 2008.

1. Tucholke, B. E. *et al.* Role of melt supply in oceanic detachment faulting and formation of megamullions. *Geology* **36**, 455–458 (2008).
2. Cann, J. R. *et al.* Corrugated slip surfaces formed at North Atlantic ridge-transform intersections. *Nature* **385**, 329–332 (1997).
3. Cannat, M. *et al.* Modes of seafloor generation at a melt-poor ultraslow-spreading ridge. *Geology* **34**, 605–608 (2006).
4. Okino, K. *et al.* Development of oceanic detachment and asymmetric spreading at the Australian–Antarctic Discordance. *Geochim. Geophys. Geosyst.* **5**, doi:10.1029/2004GC000793 (2004).
5. Tucholke, B. E., Lin, J. & Kleinrock, M. C. Megamullions and mullion structure defining oceanic metamorphic core complexes on the Mid-Atlantic Ridge. *J. Geophys. Res.* **103** (B5), 9857–9866 (1998).
6. Smith, D. K., Cann, J. R. & Escartín, J. Widespread active detachment faulting and core complex formation near 13° N on the Mid-Atlantic Ridge. *Nature* **443**, 440–444 (2006).

7. Smith, D. K. *et al.* Fault rotation and core complex formation: Significant processes in seafloor formation at slow spreading mid-ocean ridges (Mid-Atlantic Ridge, 13–15° N). *Geochim. Geophys. Geosyst.* **9**, doi:10.1029/2007GC001699 (2008).
8. Buck, W. R., Lavier, L. L. & Poliakov, A. N. B. Modes of faulting at mid-ocean ridges. *Nature* **434**, 719–723 (2005).
9. Fujiwara, T. *et al.* Crustal evolution of the Mid-Atlantic Ridge near the Fifteen-Twenty Fracture Zone in the last 5 Ma. *Geochim. Geophys. Geosyst.* **4**, doi:10.1029/2002GC000364 (2003).
10. Canales, J. P. *et al.* Seismic structure across the rift valley of the Mid-Atlantic ridge at 23° 20' N (MARK area): Implications for crustal accretion processes at slow-spreading ridges. *J. Geophys. Res.* **105** (B12), 28411–28425 (2000).
11. Smith, D. K. *et al.* Hydroacoustic monitoring of seismicity at the slow-spreading Mid-Atlantic Ridge. *Geophys. Res. Lett.* **29**, doi:10.1029/2001GL013912 (2002).
12. Bohnenstiel, D. *et al.* Aftershock sequences in the mid-ocean ridge environment: an analysis using hydroacoustic data. *Tectonophysics* **354**, 49–70 (2002).
13. deMartin, B. J. *et al.* Kinematics and geometry of active detachment faulting beneath the Trans-Atlantic Geotraverse (TAG) hydrothermal field on the Mid-Atlantic Ridge. *Geology* **35**, 711–714 (2007).
14. Smith, D. K. *et al.* Spatial and temporal distribution of seismicity along the northern Mid-Atlantic Ridge (15°–35° N). *J. Geophys. Res.* **108** (B3), doi:10.1029/2002JB001964 (2003).
15. Kong, L. S., Solomon, S. C. & Purdy, G. M. Microearthquake characteristics of a mid-ocean ridge along-axis high. *J. Geophys. Res.* **97**, 1659–1685 (1992).
16. Toomey, D. R. *et al.* Microearthquakes beneath the median valley of the Mid-Atlantic ridge near 23° N: hypocenters and focal mechanisms. *J. Geophys. Res.* **90**, 5443–5458 (1985).
17. Barclay, A. H., Toomey, D. R. & Solomon, S. C. Microearthquake characteristics and crustal Vp/Vs structure at the Mid-Atlantic Ridge, 35° N. *J. Geophys. Res.* **106** (B2), 2017–2034 (2001).
18. Wolfe, C. *et al.* Microearthquake characteristics and crustal velocity structure at 29° N of the Mid-Atlantic Ridge: The architecture of a slow-spreading segment. *J. Geophys. Res.* **100** (B12), 24449–24472 (1995).
19. Klinkhammer, G. *et al.* Hydrothermal manganese plumes in the Mid-Atlantic Ridge rift valley. *Nature* **314**, 727–731 (1985).
20. McCaig, A. *et al.* Oceanic detachment faults focus very large volumes of black smoker fluids. *Geology* **35**, 935–938 (2007).
21. Boschi, C. *et al.* Mass transfer and fluid flow during detachment faulting and development of an oceanic core complex, Atlantic Massif (MAR 30° N). *Geochim. Geophys. Geosyst.* **7**, doi:10.1029/2005GC001074 (2006).
22. Bougault, H. *et al.* Mantle heterogeneity from trace elements: MAR triple junction near 14° N. *Earth Planet. Sci. Lett.* **88**, 27–36 (1988).
23. Klein, E. & Langmuir, C. H. Local versus global variations in ocean ridge basalt composition: a reply. *J. Geophys. Res.* **94** (B4), 4241–4252 (1989).
24. Lissenberg, C. J. & Dick, H. J. B. Melt–rock reaction in the lower oceanic crust and its implications for the genesis of mid-ocean ridge basalt. *Earth Planet. Sci. Lett.* **271**, 311–325 (2008).
25. Escartín, J. *et al.* Quantifying tectonic strain and magmatic accretion at a slow spreading ridge segment, Mid-Atlantic Ridge, 29° N. *J. Geophys. Res.* **104** (B5), 10421–10437 (1999).
26. Macdonald, K. C. & Luyendyk, B. P. Deep-Tow studies of the structure of the Mid-Atlantic Ridge crest near lat 37° N. *Geol. Soc. Am. Bull.* **88**, 621–636 (1977).
27. McAllister, E. & Cann, J. R. in *Tectonic, Magmatic, Hydrothermal and Biological Segmentation of Mid-Ocean Ridges* Vol. 118 (eds MacLeod, C. J., Tyler, P. A. & Walker, C. L.) 29–48 (1996).

28. Carbotte, S. M. *et al.* New integrated data management system for Ridge2000 and MARGINS research. *EOS Trans. Am. Geophys. Un.* **85**, 553–559 (2004).
29. Müller, R. D. *et al.* Digital isochrons of the world's ocean floor. *J. Geophys. Res.* **102** (B2), 3211–3214 (1997).
30. Lehnert, K. *et al.* A global geochemical database structure for rocks. *Geochem. Geophys. Geosyst.* **1**, 1999GC000026 (2000).
31. Su, Y. J. *Mid-Ocean Ridge Basalt Trace Element Systematics: Constraints from Database Management, ICPMS Analyses, Data Compilation and Petrological Modeling* (Univ. Columbia, 2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was carried out during a 16-month visit to Harvard University and MIT by J.E., and was supported by CNRS (J.E.), NSF (D.K.S., H.S., C.L. and S.E.), WHOI (J.E., D.K.S., H.S. and J.C.), Harvard University (J.E., C.L. and S.E.), University of Leeds (J.C.) and MIT (J.E.). We thank M. Cannat and J. P. Canales for discussions, and W. R. Buck and B. Ildefonse for reviews. This is IPGP contribution 2404.

Author Contributions All authors contributed to the interpretation and analysis of the data. J.E. led the data analysis and writing of the manuscript, with contributions from all the co-authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.E. (escartin@ipgp.jussieu.fr).

METHODS

Multibeam bathymetry grids along the study area were obtained from the Marine Geoscience Data Service (<http://www.marine-geo.org>)²⁸ and from Smith *et al.*^{6,7}, and gridded at ~100 m. These data were used to define the ridge axis, either as indicated by axial volcanic ridges, or following the centre of the rift valley in the absence of these structures, as reported in the interpreted maps from Supplementary Fig. 1. The ridgeward limit of both detachment terrain and volcanic abyssal hill terrain, as defined in the text, was digitized and also reported in the maps. The ridge axis was classified as corresponding to asymmetrical accretion, symmetrical accretion, or unclassified. The relative proportions of ridge axis reported in the paper are calculated from the ridge axis geometry presented in the maps. Proportions based on ridge length may differ from the proportions apparent in Fig. 2, which corresponds to a latitudinal projection that does not take into account overlap, curvature, or gaps in ridge segmentation. Supplementary Table 1 provides relative proportions for each mode of accretion along ridge sections bound by major fracture zones, as well as their length.

Hydrophone events are from the NOAA hydroacoustic catalogue (<http://www.pmel.noaa.gov/vents/acoustics/autochart/GetPosit.html>), which extends from January 1999 to September 2003¹¹. Teleseismic events are from the NEIC teleseismic catalogue (<http://neic.usgs.gov/>) for the period 1973–2007. The digital sea-floor age map²⁹ was used to discard events occurring in lithosphere older than 3 Myr. As over 80% of the hydrophone and teleseismic events occur within 10 km of the ridge axis and corresponding to a spreading age of ~1 Myr (ref. 14), the remaining events used in the analysis presented here provide a good characterization of tectonic processes associated with accretion. Segments corresponding to symmetrical, asymmetrical and undetermined accretion were grouped into larger ridge sections with similar mode of accretion. The seismic rate (events per kilometre; Fig. 2b, c) for each ridge section was calculated using the total number of observed hydroacoustic and teleseismic events and the total length of ridge axis within each section.

Hydrothermal sites (Fig. 2; see also Supplementary Fig. 1) include active vents found so far, and inactive hydrothermal fields within the rift valley that have thus been active in geologically recent times (probably less than tens to hundreds of thousands of years). From south to north these sites are: Ashadze at 12° 58' N^{32,33}, 13° 30' N (inactive)³⁴, Logachev at 14° 45' N³⁵, Krasnov (inactive) at 16° 38' N³⁶, Snake Pit at 23° 22' N³⁷, TAG at 26° 08' N³⁸, Broken Spur at 29° 11' N³⁹ and Lost City at 30° 08' N⁴⁰. Hydrothermal plumes in the water column shown in Fig. 2 correspond to manganese anomalies generated by active hydrothermal vents within the rift valley¹⁹. In most cases the source at the sea floor of these plumes has not been directly observed.

The geochemical data set (~7,700 glass analyses) included available mid-ocean ridge basalt glass analyses from PETDB (www.petdb.org)³⁰, complemented with unpublished data from J. Cann (samples reported in ref. 41) and C. Langmuir. The sample locations were verified to retain only the analyses from basalts sampled within the rift valley (axial volcanic ridges, rift valley floor or inner rift valley walls). Major element data that could not be corrected for interlaboratory bias³¹ were then discarded, and the remaining data (~3,200 analyses) were used in the paper (see Supplementary Figs 1 and 3). The final regional data set shows differences in basalt geochemistry between asymmetrically and symmetrically accreting ridge sections, but a large overlap exists (Supplementary Fig. 3) owing to regional variations in mantle sources, among other factors. We have thus presented in Fig. 3 results for three adjacent pairs of symmetric–asymmetric spreading sections, thereby removing long-wavelength compositional variations of the mantle that cannot be attributed to differences in the mode of accretion. In Fig. 3 we also plot the liquid lines of descent for basalt to illustrate the effect of the pressure of crystallization on basalt composition. The initial composition corresponds to that of one of the most primitive basalt samples in the area. The liquid lines of descent are calculated using the program *hbasalt*⁴².

32. Davydov, M. P. *et al.* Ferromanganese deposits in the Ashadze-1 hydrothermal field (Mid-Atlantic Ridge, 12° 58' N). *Dokl. Earth Sci.* **415A**, 954–960 (2007).
33. Beltenev, V. *et al.* New hydrothermal sites at 13° N, Mid Atlantic Ridge. *InterRidge News* **14**, 14–16 (2005).
34. Beltenev, V. *et al.* A new hydrothermal field at 13° 30' N on the Mid-Atlantic Ridge. *InterRidge News* **16**, 10–11 (2007).
35. Batuyev, B. N. *et al.* Massive sulphide deposits discovered and sampled at 14° 45' N, Mid-Atlantic Ridge. *Bridge Newslett.* **6**, 6–10 (1994).
36. Beltenev, V. *et al.* A new hydrothermal field at 16° 38.4' N, 46° 28.5' W on the Mid-Atlantic Ridge. *InterRidge News* **13**, 5–6 (2004).
37. Ocean Drilling Program Leg 106 Scientific Party. Drilling the Snake Pit hydrothermal sulfite deposit on the Mid-Atlantic ridge, lat 23° 22' N. *Geology* **14**, 1004–1007 (1986).
38. Rona, P. A. *et al.* Black smokers, massive sulphides and vent biota at the Mid-Atlantic Ridge. *Nature* **321**, 33–37 (1986).
39. James, R. H., Elderfield, H. & Palmer, M. R. The chemistry of hydrothermal fluids from the Broken Spur site, 29° N Mid-Atlantic Ridge. *Geochim. Cosmochim. Acta* **59**, 651–659 (1995).
40. Kelley, D. S. *et al.* An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30° N. *Nature* **411**, 145–149 (2001).
41. Cann, J. & Smith, D. K. Evolution of volcanism and faulting in a segment of the Mid-Atlantic Ridge at 25° N. *Geochem. Geophys. Geosyst.* **6**, Q09008 (2005).
42. Langmuir, C. H. *et al.* in *Back-arc Spreading Systems: Geological, Biological, Chemical and Physical Interactions* (eds Christie, D. M. *et al.*) 87 (AGU Geophys. Monogr. Series, 2006).

Dual epithelial origin of vertebrate oral teeth

Vladimír Soukup¹, Hans-Henning Epperlein², Ivan Horáček¹ & Robert Cerný¹

The oral cavity of vertebrates is generally thought to arise as an ectodermal invagination^{1,2}. Consistent with this, oral teeth are proposed to arise exclusively from ectoderm, contributing to tooth enamel epithelium, and from neural crest derived mesenchyme, contributing to dentin and pulp^{3–5}. Yet in many vertebrate groups, teeth are not restricted only to the oral cavity^{6–9}, but extend posteriorly as pharyngeal teeth that could be derived either directly from the endodermal epithelium, or from the ectodermal epithelium that reached this location through the mouth or through the pharyngeal slits⁶. However, when the oropharyngeal membrane, which forms a sharp ecto/endodermal border¹⁰, is broken, the fate of these cells is poorly known. Here, using transgenic axolotls with a combination of fate-mapping approaches, we present reliable evidence of oral teeth derived from both the ectoderm and endoderm and, moreover, demonstrate teeth with a mixed ecto/endodermal origin. Despite the enamel epithelia having a different embryonic source, oral teeth in the axolotl display striking developmental uniformities and are otherwise identical. This suggests a dominant role for the neural crest mesenchyme over epithelia in tooth initiation and, from an evolutionary point of view, that an essential factor in teeth evolution was the odontogenic capacity of neural crest cells, regardless of possible ‘outside-in’¹¹ or ‘inside-out’¹² influx of the epithelium.

Teeth are one of the key vertebrate innovations, but their evolutionary origins are still a matter of debate. It is widely accepted that teeth initially evolved from outer skin denticles captured in the stomodeum (the odontode theory)¹¹ and modified there specifically in the context of newly developed jaws (‘outside-in’ theory). However, as there is good evidence of teeth/denticles inside the pharyngeal regions of many fossil jawless groups^{7,12}, they must have evolved with a great degree of independence from the stomodeal cavity and the jaw elements. An alternative scenario reflecting these facts has been suggested, in which oral teeth arose by the progression of ancient denticles from the endodermal pharynx towards the stomodeum (‘inside-out’ theory)¹². More recently, however, it was argued on the basis of fossil evidence that teeth may have evolved independently through a convergent evolution and, thus, are not homologous among jawed vertebrates¹³. A new, appealing hypothesis was then proposed, namely that the diversity and complexity of dentitions can be explained by combinatorial derivation of teeth from both external (ectodermal), and internal (pharyngeal) denticles⁴.

Teeth are commonly ranked among ectodermal organs⁵, although they are composite structures of dual embryonic origin. The dental mesenchyme has been shown, using a fate-mapping approach, to be derived from neural crest cells in mammals³, urodele amphibians¹⁴ (also this study; Supplementary Fig. 3) and fish¹⁵, and this is generally assumed to be the case in other vertebrates as well¹⁶. The germ-layer origin of the epithelium, however, is far less clear. Because tooth development is most completely understood in mouse embryos⁴, it is often generalized accordingly that teeth develop exclusively in the region of the oral ectoderm, which invaginates to form a stomodeum². The accepted view is that the presence of teeth in any

region is an indubitable criterion for the existence of the ectodermal germ layer in this region at some time of development¹. However, in various vertebrate lineages, so-called pharyngeal teeth, or even a second set of toothed jaws, are commonly found posterior to the stomodeum in areas that are presumably lined by endoderm rather than ectodermal epithelium^{6,8}. Convincing developmental evidence for an endodermal origin of teeth situated in the pharyngeal cavity is lacking, and uncertainties arise also from the fact that some structures situated within the pharyngeal cavity of bony fishes are apparently derived from the ectoderm (for example gills or opercular bones). Apart from the facts that the endoderm was suggested, on the basis of histology, to contribute to tooth formation in some lower vertebrates during the first half of the twentieth century^{17,18} and that such a role has been questioned even in mammals^{19,20}, our understanding of the germ-layer origin of tooth epithelia is fundamentally limited by the difficulties in distinguishing between ectoderm and endoderm during critical stages of later mouth development. Hence, after the breaking of the oropharyngeal membrane, which constitutes the border between the oral (ectodermal) and pharyngeal (endodermal) epithelia, the fate of these cells is not known, owing to a lack of reliable fate-mapping studies even for model vertebrate species like mouse, chick or zebrafish.

Urodele amphibians are an interesting group for the analysis of the germ-layer origin of teeth because the presumptive border between the oral ectoderm and endoderm is substantially more anterior than in mammals^{17,18,21}. To study mouth development and the germ-layer origin of dental tissues in details, we took advantage of recently developed transgenic axolotls²². We designed a novel experimental procedure that enables us reliably to mark the ectoderm of the entire prospective mouth area and to follow its fate during the course of development.

First we performed transplantations of four different areas of double-layered ventral epithelia using axolotl GFP-positive neurulae (Supplementary Fig. 1a; GFP, green fluorescent protein) and found conclusively that for reliable marking of the ectodermal layer of the prospective mouth, it is necessary to graft both prospective oral ectoderm and transverse neural folds (in total, $n = 113$; Fig. 1a, b). Moreover, we always ascertained that these GFP-positive grafts comprised the entire mouth area, so that no GFP-negative cells could contribute to mouth formation (Supplementary Fig. 1b).

Next we used this experimental system to trace the accurate contribution of ectodermal cells to mouth and tooth formation. In the axolotl, the epidermis in the prospective mouth region initially consists of a double-layered ectoderm but subsequently becomes reduced to a single outer layer when the inner layer bends inwards over non-ectodermal mouth tissue as an ‘ectodermal collar’^{17,18} (Fig. 1c). This oral ectodermal lining deepens (compare Supplementary Fig. 1c, d) and during later tail-bud stages contributes to prominent buds (Fig. 1d, arrowhead). However, morphologically identical budding structures also appear in the non-ectodermal area (Fig. 1d, arrow; notice the proximity to the lower-jaw cartilage, MC). Later, still before hatching, buds are easily identified as developing teeth, which are regularly distributed in both ectodermal (Fig. 1e, arrowheads) and non-ectodermal areas (Fig. 1e,

¹Department of Zoology, Charles University in Prague, Vinicna 7, 128 44 Prague, Czech Republic. ²Department of Anatomy, TU Dresden, Fetscherstrasse 74, 01307 Dresden, Germany.

arrows). To confirm the identity of these structures as tooth buds, we used an antibody directed against calbindin (Sigma), a calcium-binding protein that specifically recognizes ameloblasts²³. From sections where both GFP and immunostaining is visualized (Fig. 1f, g and Supplementary Fig. 2a, b), it is evident that the tooth primordia are developing from both ectodermal and non-ectodermal epithelia.

To substantiate our finding that in the Mexican axolotl some oral teeth develop from non-ectodermal epithelia, we invented a double-labelling approach using which cells of both oral ectoderm and foregut endoderm can be reliably marked and mapped (Fig. 2a). First, at a neurula stage, the double-layered prospective oral epithelium (from the same area as in the previous experiment) was extirpated. The exposed endodermal layer was then focally injected using the lipophilic dye DiI (Molecular Probes). Next a GFP-positive graft comprising the entire prospective oral ectoderm (as above) was transplanted orthotopically to wild-type host embryos. In this approach ($n = 91$), the entire prospective ectoderm of the oral area was marked with a green fluorescent dye and some of the foregut endodermal cells, expected to contribute to tooth buds, were labelled with a red fluorescent dye.

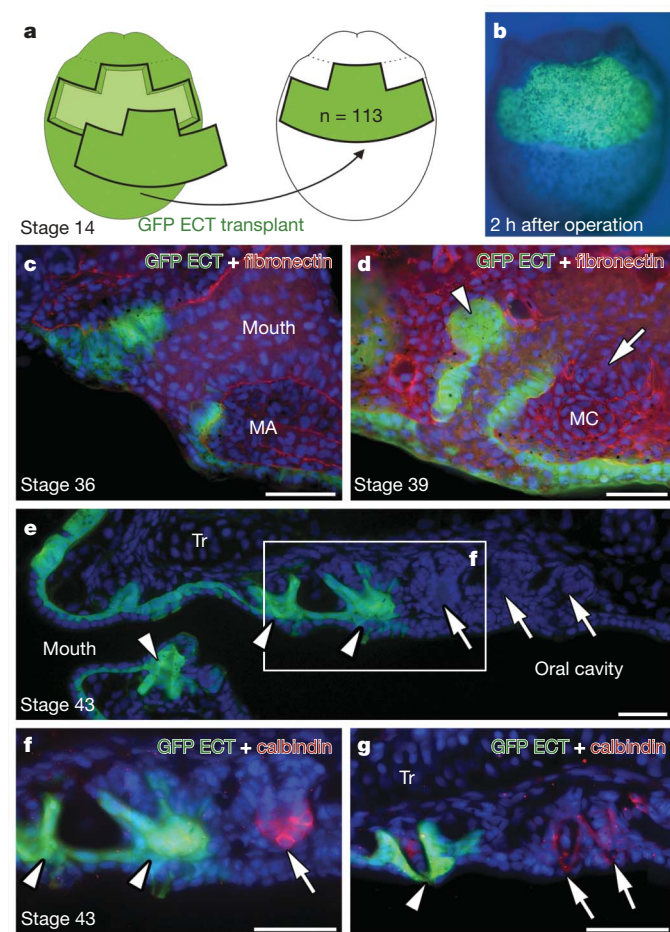


Figure 1 | Ectoderm contribution to mouth and tooth formation in the Mexican axolotl. **a**, An experimental scheme with the prospective oral ectoderm (ECT) transplanted from a GFP-positive donor to a host embryo, ventral view (END, endoderm). **b**, An embryo 2 h after operation. **c–g**, Paramedial sections, head to the left, showing a contribution of the oral ECT (green) to mouth and tooth formation. DAPI (blue) stains cell nuclei; fibronectin (red in **c**, **d**) marks cell and tissue borders. Initially the oral ECT (green) inflexes as a stomodeal collar (**c**). Then prominent tooth buds develop in ECT areas (arrowheads) as well as in non-ECT areas (arrows; **d**, **e**). Tooth buds, identified using anti-calbindin (red), develop within ECT areas (arrowheads) as well as non-ECT areas (arrows; **f**, **g**). Tr, trabecula; MA, mandibular arch; MC, Meckel's cartilage. **c**, **d**, Vibratome 100- μ m sections; **e–g**, cryostat 20- μ m sections. Scale bars, 50 μ m.

Using this double-fate-mapping approach, we obtained strong support for our previous conclusions that the axolotl possesses oral teeth with an epithelial lining of non-ectodermal origin. Specifically, dye injected into the foregut endoderm at the neurula stage was found in oral tooth germs and later in developing teeth (Fig. 2b–d, h and Supplementary Fig. 4), as well as, notably, in the epithelium situated between GFP-positive ectodermal epithelia (Fig. 2c). Moreover, alongside the contact zone between the ecto- and endodermal oral epithelia, we found tooth germs that consistently demonstrate a mixed contribution from both ecto- and endodermal cells to their enamel epithelia (Fig. 2e–g and Supplementary Fig. 5). On the basis of our combined tracing approaches, we conclude that on the upper jaw the enamel epithelia of the premaxillary/maxillary teeth are always ectodermal, whereas the enamel epithelia of the vomero-palatal teeth are derived from the ectoderm, endoderm or

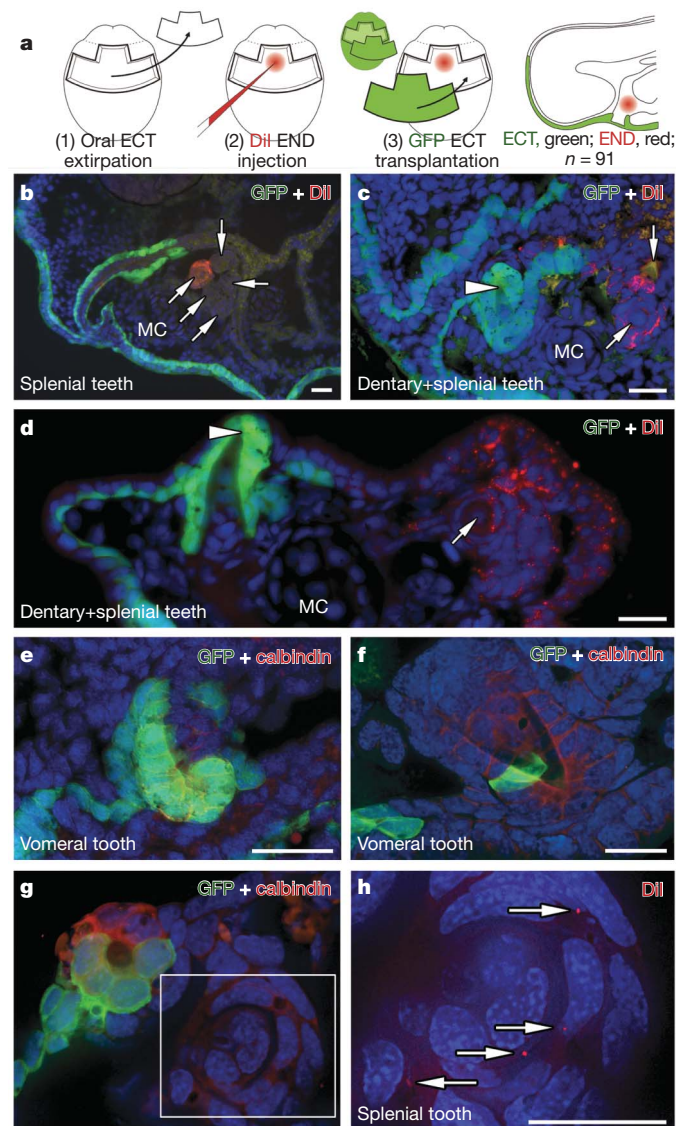


Figure 2 | Endoderm contribution to tooth formation. **a**, Sketch of double-fate-mapping experiment: following extirpation of the prospective oral ECT (1), DiI was injected into foregut END (2) and the prospective oral ECT (GFP-positive) was transplanted orthotopically (3). **b–h**, Paramedial (cryostat 20- μ m) sections, head to the left, showing a contribution of the oral ECT (green) and mouth END (red in **b–d**, **h**) to tooth formation. DAPI (blue) stains cell nuclei, calbindin (red in **e–g**) marks tooth buds. Arrows point to END teeth; arrowheads to ECT teeth. **e–g**, Details of teeth of mixed origin. **h**, A confocal image, inset in **g**, showing the END (DiI, red) contribution to splenial tooth germ. Scale bars, 25 μ m.

from a mixed source, according to their position (Fig. 3a). On the lower jaw, dentary teeth are basically ectodermal and splenial endodermal; however, there are teeth of mixed origin situated on the anterior parts of these fields (Fig. 3a).

Next a quantitative screening was performed in which all teeth were counted and their respective germ-layer origins determined at four different stages and based on 26 embryos from the double-labelling experiment (Fig. 4). This analysis revealed that of 1,137 teeth, 374 were derived from ectoderm, 598 from endoderm, and 155 were of mixed ecto/endodermal origin. We note that during the course of development, the proportion of ectoderm-derived teeth slightly increases as teeth located on the premaxillary and maxillary bones, which are purely ectodermal, develop very late. Thus, in the average embryo (analysed at stage 45, when the mouth opens and animals start to eat), of 82 teeth 29 were of ectoderm, 42 were of endoderm and 11 were of mixed epithelial origin (Supplementary Tables 1–4). Non-epithelial derivatives, such as tooth dentin and papillae, were derived from neural crest mesenchyme (from the trigeminal neural crest stream; Supplementary Fig. 3). All quantitative and statistical analyses were strongly significant (Supplementary Tables 1–4) and constitute robust support that our data are not biased by any technical problems.

Previous theories have identified the ectodermal border in the mouth as being central to tooth positioning¹¹ (Fig. 3b, upper row). However, in the Mexican axolotl, the oral ectoderm does not form a true stomodeum^{17,18}. Instead, only an inner ectodermal layer bends inwards as a stomodeal collar over the dense endodermal rod, which blocks the prospective mouth at early stages of development (Fig. 3b,

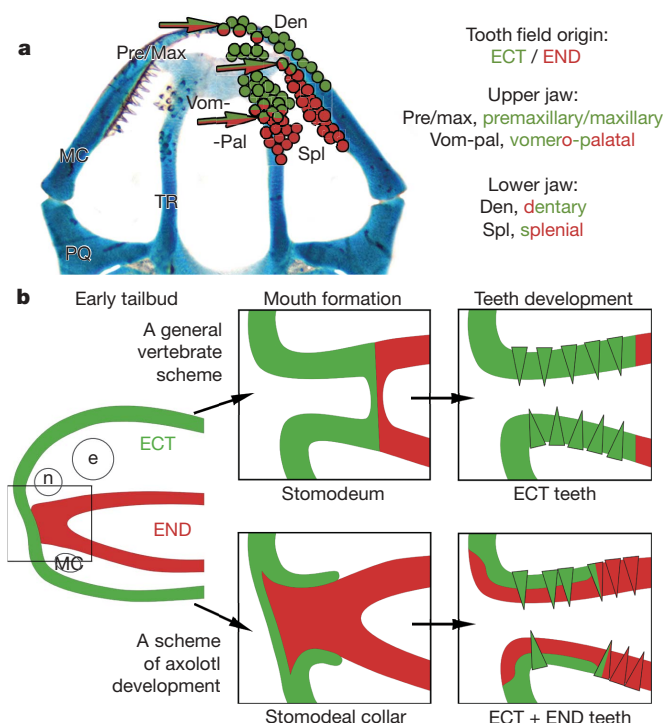


Figure 3 | Germ-layer origin and morphogenesis of teeth of the Mexican axolotl. **a**, A sketch of the germ-layer origin of teeth in the Mexican axolotl. ECT teeth, green; END teeth, red; teeth of mixed origin, red-green (the colouring in the key is a qualitative guide to the ratio of the components in each tooth field). Cartilages visualized using alcian blue. **b**, Comparative developmental morphogenesis of the mouth region and the germ-layer origin of teeth of vertebrate (upper row) and an axolotl (lower row) embryo. In the majority of vertebrates, the mouth develops from a stomodeum with teeth distributed in invaginated ECT. In contrast, in urodeles the mouth develops from a stomodeal collar with an oral epithelium either of a dual origin, with teeth of ECT or END, or of a mixed origin. PQ, palatoquadrate; TR, trabecula cranii; n, nose; e, eye.

lower row). Because of this positioning, the collar cells develop into the basal cells, and the outer cells of the endodermal rod develop into the apical cells of the oral epithelium during the course of mouth opening (Fig. 1e and Supplementary Fig. 2a–d; summarized in Fig. 3b, lower row). The endodermal cells of the mouth, as part of the epithelial lining, are consequently found also on the outer surface of the mouth (Fig. 1e and Supplementary Fig. 2b–d, arrowheads; summarized in Fig. 3b, lower row). Thus, in the axolotl, the posterior part of the oral cavity is lined with the endodermal epithelium, whereas the anterior part is lined with an epithelium of double origin (Fig. 3b and Supplementary Fig. 2b). This provides reliable documentation of an oral endodermal epithelial lining that reaches outside the mouth, and, also, of an oral epithelium originating from two germ layers. A considerable number of reports on mouth development have been published, but, as underlined by our results, there is still a need for detailed fate-mapping approaches in studies of dynamic interactions of cells and tissues derived from different germ layers.

Progressing from recent vivid discussions on the subject^{4,6,8,24}, our data present reliable evidence of oral teeth of endodermal origin in vertebrates. We speculate that oral teeth of endodermal origin might form in all animals with oral endoderm, that is, in urodele and probably also lungfish species, and maybe even in some frogs²⁵, where the mouth develops from a structure similar to the stomodeal collar. However, as a possible interdigitation of cells from both epithelial tissue layers during mouth formation has not been fate-mapped for any vertebrate species, and some reports indicate that foregut endoderm may stretch more to

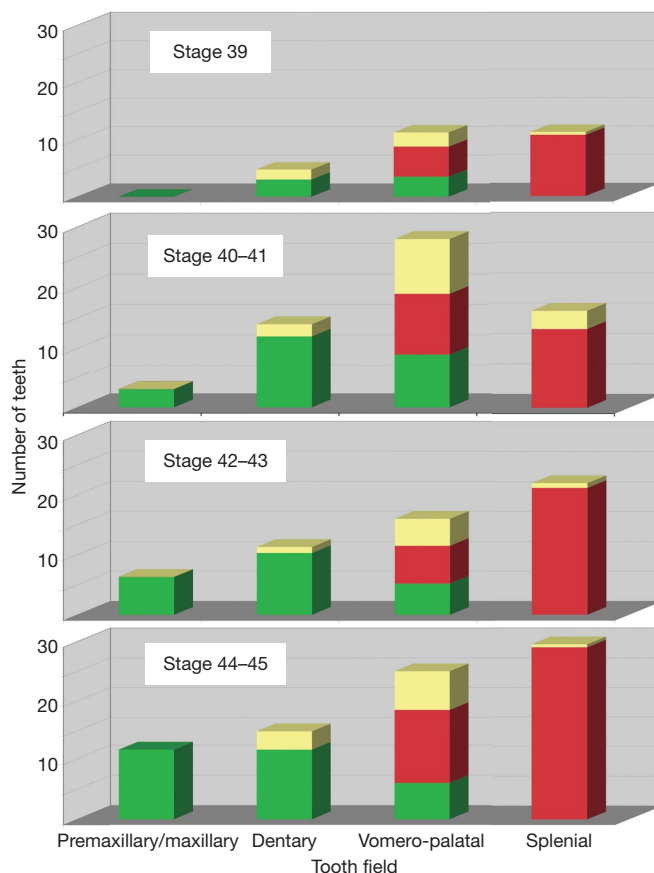


Figure 4 | A quantitative screening of teeth from different tooth fields with respect to their germ-layer origin. $n = 1,137$ teeth: 374 ECT (green), 598 END (red), 155 ECT–END mixture (yellow). Data based on 26 animals, visualized at four stages: four animals at stage 39, 14 animals at stage 40–41, six animals at stage 42–43, two animals at stage 44–45. For details and original data, see Supplementary Tables 1–4.

the anterior than hitherto believed^{8,19–21}, we speculate that oral teeth of endodermal origin might present a more common feature in vertebrate oral development than previously assumed.

Whereas the classical 'outside-in' theory implies that teeth were initially derived from the oral ectodermal layer¹¹, the 'inside-out' theory strongly suggests that they were derived from the endodermal layer¹², and this derivation is believed to impart differences to denticles, teeth or dentition in terms of shape and complexity^{4,12,24}. However, the dual origin of enamel epithelia in otherwise morphologically identical axolotl oral tooth primordia (as regards complexity, shape, position, timing and morpho-differentiation of teeth), together with studies illustrating deep shared molecular similarities between oral (supposedly ectodermal) and pharyngeal (supposedly endodermal) teeth^{15,26,27} imply that 'ectodermal' and 'endodermal' teeth do not differ essentially. It is beyond the scope of this study to identify the plesiomorphic germ-layer origin of tooth epithelium. However, our results clearly demonstrate that the germ-layer origin of epithelium into which the mesenchyme cells come into contact does not affect the final product of the odontogenic cascade. Mesenchyme cells can thus apparently interact with a host of epithelial cells, forming teeth/denticles when in the stomodeum, in the pharyngeal cavity or on the skin surface. All this suggests that the major agent of dental development is the neural crest mesenchyme rather than the epithelium, the role of which in tooth patterning^{5,12,28} and even in tooth initiation²⁹ may be less fundamental than commonly believed. It therefore seems most likely that all teeth of extant vertebrates—or, more precisely, the developmental machinery producing them—have evolved only once, somewhere in the oropharynx, driven by a neural crest signal.

METHODS SUMMARY

Embryos. Embryos of the Mexican axolotl (*Ambystoma mexicanum*) were obtained, reared and staged as previously described³⁰. GFP embryos were obtained from the Max-Planck-Institute of Molecular Cell Biology and Genetics in Dresden, Germany, and were developed in the laboratory of E. Tanaka²².

Operations and injections. GFP ectodermal transplantations were performed as sketched in Fig. 1a ($n = 113$). At first, however, transplantations of four different areas of double-layered ventral epithelia were performed (Supplementary Fig. 1a) to define the entire ectodermal layer of the prospective mouth.

The double-labelling approach by which cells of both oral ectoderm and foregut endoderm were marked and mapped (as sketched in Fig. 2a ($n = 91$)) includes extirpation of the double-layered prospective oral epithelia (from the same area as in the previous experiment), focal injection of a CellTracker CM-DiI (Molecular Probes) into the exposed endodermal layer and, lastly, the orthotopic transplantation of a GFP-positive prospective oral ectoderm (as above) into wild-type host embryos.

Sectioning and immunostaining. Axolotl embryos were anaesthetized using MS-222 (Sigma), fixed in 4% paraformaldehyde in phosphate buffered saline and sectioned using a Vibratome 1000 sectioning system (Ted Pella) or a CM3050 cryostat (Leica). Sections were counterstained using anti-fibronectin antibody (Dako) to visualize tissue borders, with 4,6-diamidino-2-phenylindole (DAPI) to mark cell nuclei, or with anti-calbindin antibody (Sigma), which specifically recognizes ameloblasts²³.

Image acquisition. Separate fluorescence images were captured using an Olympus BX51 microscope with a SPOT RT camera, or the Olympus Cell^R IX81 with a Hamamatsu Photonics Orca camera, merged and optimized using Spot and Adobe Photoshop software.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 January; accepted 29 July 2008.

Published online 14 September 2008.

1. Romer, A. S. *The Vertebrate Body* (Saunders, 1962).
2. Kardong, K. V. *Vertebrates. Comparative Anatomy, Function, Evolution* (Brown, 1995).
3. Chai, Y. et al. Fate of the mammalian cranial neural crest during tooth and mandibular morphogenesis. *Development* **127**, 1671–1679 (2000).
4. Tucker, A. & Sharpe, P. The cutting-edge of mammalian development: how the embryo makes teeth. *Nature Rev. Genet.* **5**, 499–508 (2004).
5. Pispis, J. & Thesleff, I. Mechanisms of ectodermal organogenesis. *Dev. Biol.* **262**, 195–205 (2003).

6. Stock, D. W. The genetic basis of modularity in the development and evolution of the vertebrate dentition. *Phil. Trans. R. Soc. Lond. B* **356**, 1633–1653 (2001).
7. Janvier, P. *Early Vertebrates* (Oxford Univ. Press, 1996).
8. Huysseune, A., Van der Heyden, C., Verreijdt, L., Wautier, K. & Van Damme, N. Fish dentitions as paradigms for odontogenic questions. *Connect. Tissue Res.* **43**, 98–102 (2002).
9. Sire, J. Y., Davit-Beal, T., Delgado, S., Van Der Heyden, C. & Huysseune, A. First-generation teeth in nonmammalian lineages: evidence for a conserved ancestral character? *Microsc. Res. Tech.* **59**, 408–434 (2002).
10. Dickinson, A. J. & Sive, H. Development of the primary mouth in *Xenopus laevis*. *Dev. Biol.* **295**, 700–713 (2006).
11. Reif, W.-E. Evolution of dermal skeleton and dentition in vertebrates: the odontode-regulation theory. *Evol. Biol.* **15**, 287–368 (1982).
12. Smith, M. M. & Coates, M. I. in *Major Events in Early Vertebrate Evolution* (ed. Ahlberg, P. E.) 223–240 (Taylor and Francis, 2001).
13. Smith, M. M. & Johanson, Z. Separate evolutionary origins of teeth from evidence in fossil jawed vertebrates. *Science* **299**, 1235–1236 (2003).
14. Chibon, P. Analyse expérimentale de la régionalisation et des capacités morphogénétiques de la crete neurale chez l'Amphibien Urodèle *Pleurodeles waltlii* Michah. *Mem. Soc. Zool. Fr.* **36**, 1–107 (1966).
15. Jackman, W. R., Draper, B. W. & Stock, D. W. Fgf signaling is required for zebrafish tooth development. *Dev. Biol.* **274**, 139–157 (2004).
16. Hall, B. K. *The Neural Crest in Development and Evolution* (Springer, 1999).
17. de Beer, G. R. The differentiation of neural crest cells into visceral cartilages and odontoblasts in *Amblystoma*, and re-examination of the germ-layer theory. *Proc. R. Soc. Lond. B* **134**, 377–398 (1947).
18. Adams, A. E. An experimental study of the development of the mouth in the amphibian embryo. *J. Exp. Zool.* **40**, 311–379 (1924).
19. Imai, H., Osumi, N. & Eto, K. Contribution of foregut endoderm to tooth initiation of mandibular incisor in rat embryos. *Eur. J. Oral Sci.* **106** (Suppl. 1), 19–23 (1998).
20. Sharpe, P. T. Homeobox genes and orofacial development. *Connect. Tissue Res.* **32**, 17–25 (1995).
21. Barlow, L. A. & Northcutt, R. G. Embryonic origin of amphibian taste buds. *Dev. Biol.* **169**, 273–285 (1995).
22. Sobkow, L., Epperlein, H. H., Herklotz, S., Straube, W. L. & Tanaka, E. M. A germline GFP transgenic axolotl and its use to track cell fate: dual origin of the fin mesenchyme during development and the fate of blood cells during regeneration. *Dev. Biol.* **290**, 386–397 (2006).
23. Barlow, L. A. & Northcutt, R. G. Taste buds develop autonomously from endoderm without induction by cephalic neural crest or paraxial mesoderm. *Development* **124**, 949–957 (1997).
24. Smith, M. M. & Coates, M. I. Evolutionary origins of the vertebrate dentition: phylogenetic patterns and developmental evolution. *Eur. J. Oral Sci.* **106** (Suppl. 1), 482–500 (1998).
25. Reiss, J. O. Early development of chondrocranium in the tailed frog *Ascaphus truei* (Amphibia: Anura): implications for anuran palatoquadrate homologies. *J. Morphol.* **231**, 63–100 (1997).
26. Debais-Thibaud, M. et al. Development of oral and pharyngeal teeth in the medaka (*Oryzias latipes*): comparison of morphology and expression of *eve1* gene. *J. Exp. Zool. B Mol. Dev. Evol.* **308**, 693–708 (2007).
27. Fraser, G. J., Graham, A. & Smith, M. M. Conserved deployment of genes during odontogenesis across osteichthyan. *Proc. R. Soc. Lond. B* **271**, 2311–2317 (2004).
28. Smith, M. M. Vertebrate dentitions at the origin of jaws: when and how pattern evolved. *Evol. Dev.* **5**, 394–413 (2003).
29. Mitsiadis, T. A., Caton, J. & Cobourne, M. Waking-up the sleeping beauty: recovery of the ancestral bird odontome program. *J. Exp. Zool. B Mol. Dev. Evol.* **306**, 227–233 (2006).
30. Cerny, R. et al. Combined intrinsic and extrinsic influences pattern cranial neural crest migration and pharyngeal arch morphogenesis in axolotl. *Dev. Biol.* **266**, 252–269 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We wish to thank E. Tanaka and H. Andreas for supplying us with GFP axolotl embryos, L. Mchedlishvili for technical advice and O. Sebesta for help with confocal microscopy. We are grateful to M. Bronner-Fraser, R. Ericsson, D. Meulemans Medeiros, L. Olsson, P. D. Polly, D. W. Stock and A. Tucker for comments on the manuscript. Grants from the Ministry of Youth, Education and Sport of the Czech Republic (MSMT 0021620828), SMWK (Dresden) and COST Action B23 'Oral facial development and regeneration' are gratefully acknowledged.

Author Contributions Tissue grafting was carried out in the laboratory of H.-H.E., who also helped with initial experiments. V.S. made transplantations and performed sectioning and image analyses as his Master's thesis under the supervision of R.C., who also designed the initial experiments and wrote the manuscript. R.C. and I.H. planned the study and interpreted the results, I.H. made statistical analyses and helped in writing the manuscript, and all authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to R.C. (cerny8@natur.cuni.cz).

METHODS

Embryos. Embryos of the Mexican axolotl (*A. mexicanum*) were obtained, reared and staged as previously described^{30,31}. GFP embryos were obtained from the Max-Planck-Institute of Molecular Cell Biology and Genetics in Dresden, Germany, and were spawned from a β -actin promoter-driven GFP germ-line transgenic animal that had been produced by plasmid injection²². Embryos were kept in tap water, and before being used for transplantations and injections, embryos were washed thoroughly with tap water and sterile Steinberg solution containing antibiotics (Antibiotic-Antimycotic, Gibco) and then decapsulated manually.

Operations and injections. GFP ectodermal transplantations were performed as sketched in Fig. 1a ($n = 113$). Operations were performed under sterile conditions using tungsten needles in an agar dish containing 1 M Steinberg solution plus antibiotics. We designed an experimental procedure that enabled us to mark the ectoderm of the entire prospective mouth area reliably and to follow its fate during the course of development. First we performed transplantations of four different areas of double-layered ventral epithelia from GFP-positive to host neurulae (Supplementary Fig. 1a; numbers of animals used for each operation are indicated there) and found conclusively that for reliable marking of the ectodermal layer of the prospective mouth it is necessary to graft both prospective oral ectoderm and a transverse neural fold (Fig. 1a, b). Using this type of transplantation, we were able to follow the fate of the entire ectodermal layer that translocates into the mouth, and, therefore, in this way all ectoderm-derived teeth became GFP-positive. We always ascertained, however, that these GFP grafts comprised the entire mouth area, so that no GFP-negative cells could contribute to mouth formation (Supplementary Fig. 1b).

Next we invented a double-labelling approach by which cells of both oral ectoderm and foregut endoderm can be reliably marked and mapped (Fig. 2a). First, at a neurula stage, double-layered prospective oral epithelia (from the same area as in the previous experiment) were extirpated. The exposed endodermal layer was then focally injected using the lipophilic dye DiI (Molecular Probes), dissolved in absolute ethanol to a concentration of 1 mg ml^{-1} and further diluted in nine parts of 10% sucrose in water just before injection. Then a graft from a GFP-positive neurula comprising the entire prospective oral ectoderm (as above) was transplanted orthotopically to wild-type host embryos. In this approach ($n = 91$), the entire prospective ectoderm of the oral area was marked green (GFP) and some of foregut endodermal cells, expected to contribute to tooth buds, were labelled red (DiI).

Neural crest transplantations. Trigeminal neural crest cells were transplanted from GFP-positive to wild-type embryos at the neurula stage as described in detail elsewhere³².

31. Epperlein, H. H., Meulemans, D., Bronner-Fraser, M., Steinbeisser, H. & Selleck, M. A. Analysis of cranial neural crest migratory pathways in axolotl using cell markers and transplantation. *Development* **127**, 2751–2761 (2000).
32. Cerny, R. *et al.* Developmental origins and evolution of jaws: new interpretation of “maxillary” and “mandibular”. *Dev. Biol.* **276**, 225–236 (2004).

The genome of the simian and human malaria parasite *Plasmodium knowlesi*

A. Pain^{1*}, U. Böhme^{1*}, A. E. Berry^{1*}, K. Mungall¹, R. D. Finn¹, A. P. Jackson¹, T. Mourier², J. Mistry¹, E. M. Pasini³, M. A. Aslett¹, S. Balasubramaniam¹, K. Borgwardt⁴, K. Brooks¹, C. Carret¹, T. J. Carver¹, I. Cherevach¹, T. Chillingworth¹, T. G. Clark^{1,5}, M. R. Galinski⁶, N. Hall⁷, D. Harper¹, D. Harris¹, H. Hauser¹, A. Ivens¹, C. S. Janssen⁸, T. Keane¹, N. Larke¹, S. Lapp⁶, M. Marti⁹, S. Moule¹, I. M. Meyer¹⁰, D. Ormond¹, N. Peters¹, M. Sanders¹, S. Sanders¹, T. J. Sargeant^{11,12}, M. Simmonds¹, F. Smith¹, R. Squares¹, S. Thurston¹, A. R. Tivey¹, D. Walker¹, B. White¹, E. Zuiderwijk¹, C. Churcher¹, M. A. Quail¹, A. F. Cowman¹¹, C. M. R. Turner⁸, M. A. Rajandream¹, C. H. M. Kocken³, A. W. Thomas³, C. I. Newbold^{1,13}, B. G. Barrell¹ & M. Berriman¹

Plasmodium knowlesi is an intracellular malaria parasite whose natural vertebrate host is *Macaca fascicularis* (the 'kra' monkey); however, it is now increasingly recognized as a significant cause of human malaria, particularly in southeast Asia^{1,2}. *Plasmodium knowlesi* was the first malaria parasite species in which antigenic variation was demonstrated³, and it has a close phylogenetic relationship to *Plasmodium vivax*⁴, the second most important species of human malaria parasite (reviewed in ref. 4). Despite their relatedness, there are important phenotypic differences between them, such as host blood cell preference, absence of a dormant liver stage or 'hypnozoite' in *P. knowlesi*, and length of the asexual cycle (reviewed in ref. 4). Here we present an analysis of the *P. knowlesi* (H strain, Pk1(A+) clone⁵) nuclear genome sequence. This is the first monkey malaria parasite genome to be described, and it provides an opportunity for comparison with the recently completed *P. vivax* genome⁴ and other sequenced *Plasmodium* genomes^{6–8}. In contrast to other *Plasmodium* genomes, putative variant antigen families are dispersed throughout the genome and are associated with intrachromosomal telomere repeats. One of these families, the KIRs⁹, contains sequences that collectively match over one-half of the host CD99 extracellular domain, which may represent an unusual form of molecular mimicry.

The *P. knowlesi* genome sequence was produced by whole-genome shotgun sequencing to eightfold coverage, with targeted gap closure and finishing (Supplementary Table 1). The 23.5-megabase (Mb) nuclear genome is composed of 14 chromosomes and contains the expected complement of non-coding RNA (ncRNA) genes with known function (Supplementary Table 2) and a large number of novel structured ncRNA candidate genes (Supplementary Figs 1–5 and Supplementary Tables 3 and 4). The presumed centromeres are similar to those found in other *Plasmodium* species^{4,6}, and are positionally conserved within regions sharing synteny with *P. vivax* (see Fig. 1 of ref. 4). The overall G+C base composition is 37.5%. A total of 5,188 protein-encoding genes were identified, which is

slightly lower than the predicted proteome size of *P. falciparum* and *P. vivax*^{4,6}.

Unusually for *Plasmodium* species, (G+C)-rich repeat regions containing intrachromosomal telomeric sequences (ITSs, containing the heptad sequence GGGTT[T/C]A) are found at multiple internal sites in the *P. knowlesi* chromosomes, arrayed tandemly or as components of larger repeat units (Fig. 1). These sequences appear infrequently in *P. vivax* and *P. falciparum* at internal chromosome sites (Supplementary Figs 6 and 7). In the protozoan parasite *Trypanosoma brucei*¹⁰, ITSs may be the templates for recombination events that result in gene conversion among variant antigen VSG genes¹¹. In mammalian genomes¹², ITSs are common and may represent the 'scars' of double-stranded DNA break repair¹². Alternatively, ITSs may have a role in transcriptional control.

For approximately 80% (4,156 out of 5,185) of predicted genes in *P. knowlesi*, orthologues could be identified in both *P. falciparum* and *P. vivax* (for details, see ref. 4). The *P. knowlesi*-specific variant antigen gene families, *SICAvar* genes¹³ and *kir* genes⁹, form the largest groups of *P. knowlesi*-specific expansions (Supplementary Tables 5 and 6). Five distinct gene families of unknown function, with 4–15 paralogous members, are unique to *P. knowlesi* (referred to as Pk-fam-a to Pk-fam-e in Supplementary Table 7). Pk-fam-a and Pk-fam-b each have more than nine paralogous members (Supplementary Fig. 8), which have a two-exon gene structure with a signal peptide, a carboxy-terminal transmembrane region, but lack typical export motifs^{14,15}. Members of the protein family Pk-fam-c and Pk-fam-e represent two new families with putative protein export signals (Supplementary Fig. 8 and Supplementary Table 8).

A comparison of Pfam domains¹⁶ between the predicted proteomes of *P. knowlesi*, *P. vivax* and *P. falciparum* (Supplementary Table 9, Supplementary Information) revealed major differences in domains that distinguish species-specific protein families involved in antigenic variation. The remainder of the proteome was relatively conserved albeit with some interesting copy number variations of a

¹Wellcome Trust Sanger Institute, Genome Campus, Hinxton, Cambridgeshire CB10 1SA, UK. ²Ancient DNA and Evolution Group, Department of Biology, University of Copenhagen, DK-2100 Copenhagen, Denmark. ³Department of Parasitology, Biomedical Primate Research Centre, PO Box 3306, 2280 GH, Rijswijk, The Netherlands. ⁴Machine Learning Group, Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK. ⁵Wellcome Trust Centre for Human Genetic, University of Oxford, Roosevelt Drive, Oxford OX3 9BN, UK. ⁶Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, 954 Gatewood Road, Atlanta, Georgia 30329, USA. ⁷School of Biological Sciences, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK. ⁸Institute of Biomedical and Life Sciences and Wellcome Centre for Molecular Parasitology, University of Glasgow, 120 University Place, Glasgow G12 8TA, UK. ⁹Department of Immunology and Infectious Diseases, Harvard School of Public Health, 677 Huntington Avenue, Boston, Massachusetts 02115, USA. ¹⁰UBC Bioinformatics Centre and Department of Computer Science, University of British Columbia and Department of Medical Genetics, 2366 Main Mall, British Columbia, Vancouver V6T 1Z4, Canada. ¹¹The Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria 3050, Australia. ¹²The Department of Medical Biology, The University of Melbourne, Parkville, Victoria 3010, Australia. ¹³The Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DS, UK.

*These authors contributed equally to this work.

few key housekeeping enzymes (Supplementary Fig. 9 and Supplementary Table 9).

In other *Plasmodium* genomes sequenced so far, variant gene families involved in antigenic variation (Supplementary Figs 6 and 7) are typically arranged in the subtelomeres, and only a few members of these families have hitherto been found at intrachromosomal sites. Notably, the *P. knowlesi* genome sequence has revealed that the major variant gene families (that is, *SICAvar*¹³ and *kir*⁹) are randomly distributed across all 14 chromosomes (Fig. 1) and often co-localize with ITS-containing repeats (Supplementary Information). Although all of the telomeres were not fully assembled, we know that in the case of chromosome 7, *P. knowlesi* and *P. vivax* have atypical gene content—the subtelomere encodes proteins associated with merozoite invasion (for example, MAEBL and members of the reticulocyte-binding-like (RBL) family) (Supplementary Fig. 10).

Variant SICA (schizont-infected cell agglutination) antigens on the surface of infected red blood cells⁵ are associated with parasite virulence¹⁷ and are encoded by the *SICAvar* gene family¹³—the largest variant antigen gene family in *P. knowlesi*. Switching of variant types underlies the establishment of a chronic infection in the vertebrate host, a process that is essential in all species, to ensure

mosquito transmission and the completion of the life cycle. Full-length *SICAvar* genes have 3–14 exons (Supplementary Table 5 and Supplementary Fig. 11), resulting in a range of sizes for the predicted proteins of 53–247 kDa. Although many of the *SICAvar* genes are present only as fragments, we estimate that there are up to 107 members in the H strain of *P. knowlesi* based on the number of conserved final exons.

Twenty-nine predicted *SICAvar* genes have complete gene structures and were divided into two subtypes (Fig. 2). The type I *SICAvar* genes with 7–14 exons predominate, with a few containing unusually long introns (Fig. 2). The type II subgroup represents small *SICAvar* genes with 3–4 exon structures. Unusually large introns (5.8–13.6 kb) are a unique feature of *SICAvar* genes and have not previously been seen in any other sequenced apicomplexan gene (Fig. 2).

SICA antigens have a modular structure (Fig. 3, Supplementary Fig. 12) comprising a variable number of highly diverged cysteine-rich domains (CRDs) encoded by multiple exons, a transmembrane domain and a cytoplasmic domain. A high level of sequence diversity was observed, with the exception of the 3' terminal exon¹³. We investigated the domain organization of the CRDs using profile hidden Markov models (HMMs; Fig. 3 and Supplementary Fig. 13). The

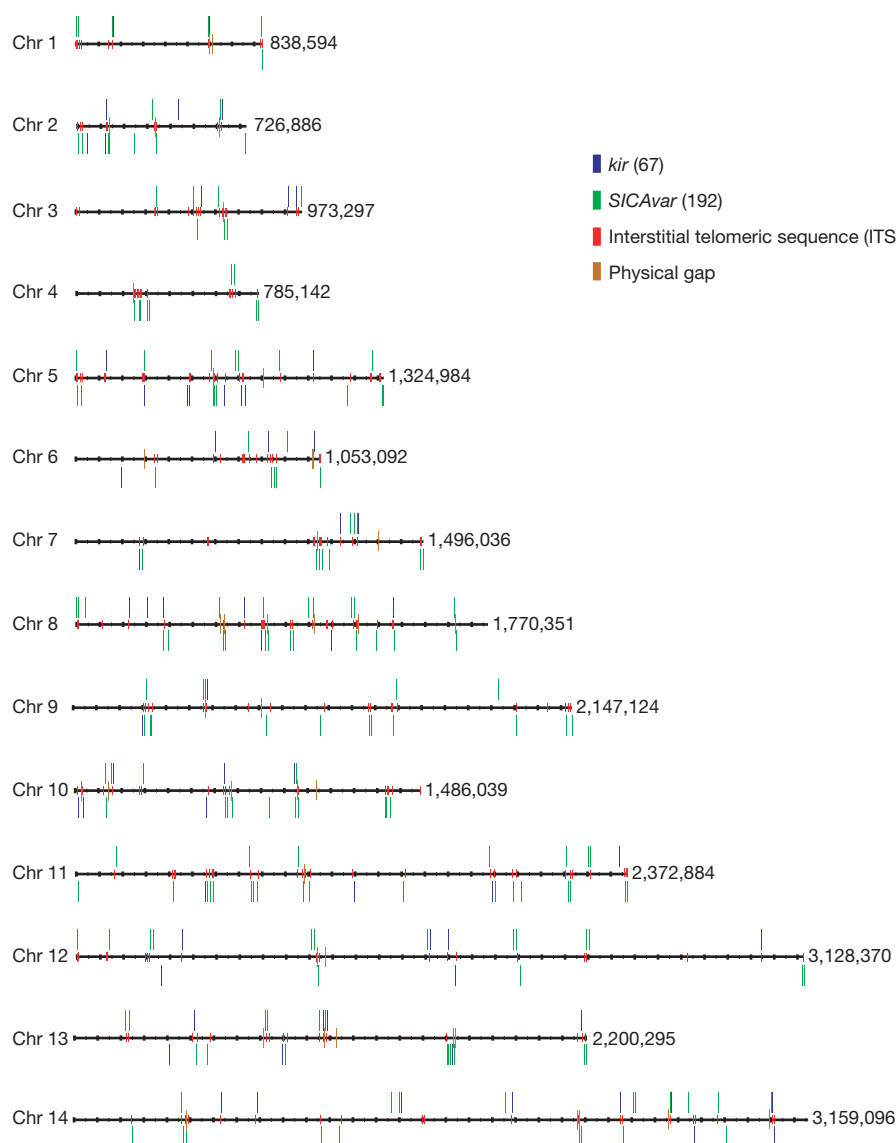


Figure 1 | Distribution of *SICAvar* genes, *kir* genes and telomere-like repeats on chromosomes 1 to 14 of *P. knowlesi* (H strain). The positions of *kir* (shown in blue) and *SICAvar* (green) genes and gene fragments are shown on all 14 chromosomes. Interstitial telomeric sequences

(GGGTT[T/C]A) are found surrounding *kir* and *SICAvar* genes (shown in red). The values along the right of each chromosome indicate the total sequence length in base pairs.

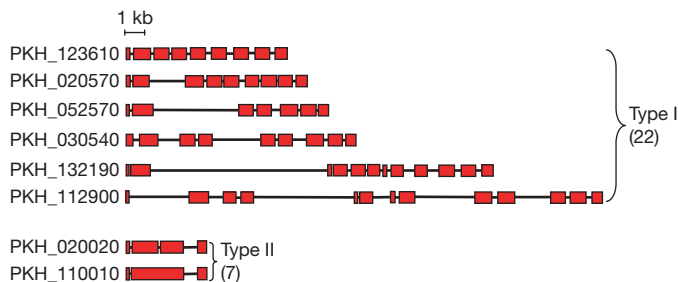


Figure 2 | Structural organization of complete (full length) *SICAvAr* genes in *P. knowlesi* (H strain). Schematic view of the exon structure of type I and type II *SICAvAr* genes. Exons are shown as red boxes with introns as joining lines.

full-length SICA proteins contain a distinct five-cysteine CRD (termed SICA- α) at the amino terminus, which occurs once or twice and may have a stabilizing role analogous to the cysteine-rich N-terminal capping motifs of extracellular leucine-rich repeat proteins¹⁸. There are 1–8 CRDs (referred to as SICA- β) with 7–10 conserved cysteine residues. The transmembrane domain and a conserved domain follow at the C terminus (termed SICA_C in Supplementary Figs 12 and 13).

Although *P. knowlesi* and *P. falciparum* are phylogenetically distant, the SICA and *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) variant antigens share many fundamental biological characteristics (reviewed in ref. 19). Common regulatory mechanisms involving post-transcriptional gene silencing have been proposed between the *var* gene family in *P. falciparum* and the *SICAvAr* family in *P. knowlesi*¹⁹. We have identified conserved sequence motifs between the single *var* intron and *SICAvAr* introns (Supplementary Figs 14–18) in the region thought to be the origin of a ncRNA transcript involved in the silencing of *var* genes²⁰, indicating possible commonality in regulatory mechanisms.

We searched for evidence of gene conversion within the *SICAvAr* family, using the predicted sequences of 20 type I full-length *SICAvAr* genes (Supplementary Information). It is clear that exon shuffling has an important role in *SICAvAr* evolution¹³. The low-complexity repeat regions found within introns might facilitate recombination

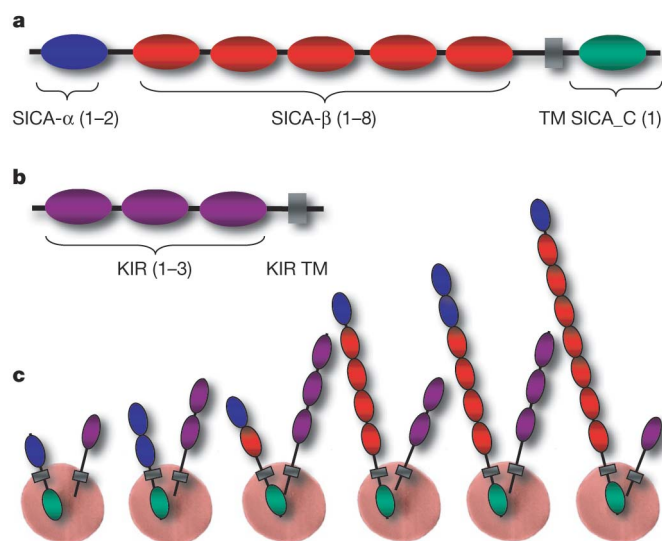


Figure 3 | Domain organization of complete (full-length) SICA and KIR proteins in *P. knowlesi* (H strain). **a**, Domain organization of full-length SICA proteins. The number of different domains (SICA- α , SICA- β and SICA_C) is shown in parentheses. TM, transmembrane. **b**, Domain organization of full-length KIR proteins. **c**, Examples of an infected erythrocyte showing SICA and KIR proteins anchored to the surface in different combinations.

through misalignment during mitosis; this could explain the presence of *SICAvAr* fragments found throughout the genome and/or *SICAvAr* gene models with partial intron/exon structures. These comprise whole, and apparently intact, exons that might provide a reservoir for diversification analogous to that seen with VSG genes in *Trypanosoma brucei*¹¹ (Supplementary Information).

Kirs represent the second largest variant gene family. They encode predicted proteins of 36–97 kDa that are hypothesized to be expressed at the surface of infected erythrocytes and undergo antigenic variation⁹. There are 68 predicted *kir* genes, 4 of which have incomplete structures (Supplementary Table 6). They were divided into four types depending on the number of exons (Supplementary Fig. 19). Most (58 out of 64) *kir* genes belong to types I and II. The domain organization of all predicted KIR proteins was also determined using profile HMMs (Fig. 3 and Supplementary Fig. 20). They contain 1–3 domains, followed by a transmembrane domain at the C terminus (referred to as KIR TM in Supplementary Fig. 20). A BLAST analysis of KIR proteins revealed stretches of up to 36 amino acids within the predicted extracellular domain that have 100% identity to host proteins, the most striking of which is to CD99. These matches were evident in several KIR proteins. Interestingly, different family members contain matches to different regions of CD99, such that together, they represent over one-half of the CD99 extracellular domain (Fig. 4). Tests were performed to assess the possibility that such matches could occur by chance (Supplementary Table 10). We have compared the sequences to *Macaca mulatta*, African green monkey and human. The matches exclude conserved cysteine regions and the degree of sequence identity decreases noticeably as the evolutionary distance to the natural host increases (Fig. 4 and Supplementary Table 10). CD99 has a critical role as an immunoregulatory molecule in T-cell function (see <http://www.ncbi.nlm.nih.gov/omim/>). These exact matches may interfere with recognition of parasitized erythrocytes by the host immune system or act as CD99 analogues that interfere by competing with T cells for CD99 partner molecules.

We undertook a more systematic search for other such instances of parasite proteins containing extensive stretches of identical host sequences, using the PMATCH algorithm (Supplementary Information). Unsurprisingly, a large number of matches to highly conserved housekeeping genes were observed, but in addition regions of perfect identity to another host protein (known as AHNAK, see <http://www.ncbi.nlm.nih.gov/omim/>) were detected in two KIRs and one SICA-like protein (Supplementary Fig. 21 and Supplementary Table 10). Analogous searches using the predicted exported protein repertoires (exportome) of *P. vivax* and *P. falciparum* found no such matches to host proteins (Supplementary Table 11). The identity to host proteins is maintained at the amino acid sequence rather than DNA sequence level (data not shown).

Acquisition of host proteins, and thus the ability to mimic their function, has been observed in many bacterial and viral pathogens²¹. In parasitic protozoa there are known cases where stretches of amino acids present on a parasite-encoded cell surface protein match perfectly to regions of host proteins²². However, in all such cases, the matches correspond to a common amino acid repeat that is shared between them^{22–24}. Malaria parasites are known to have a potential immunomodulatory role either by secreting functional homologues of host molecules or by binding to host antigen-presenting cells^{25,26}. This is the first observation of its kind in a malaria protein that shows acquisition of host peptide sequences that are likely to be on the infected cell surface and thus may interact with the host. The mechanism by which these host sequences have arisen remains to be clarified. Possible explanations include convergent evolution or horizontal transfer followed by gene degeneration events.

During the intraerythrocytic life cycle, malaria parasites significantly remodel the erythrocyte by exporting numerous proteins^{14,15}. This depends on a short motif, termed the plasmodium export element (PEXEL) or vacuolar transport signal (VTS), which is present in over 300 *P. falciparum* proteins and is common to all *Plasmodium* species

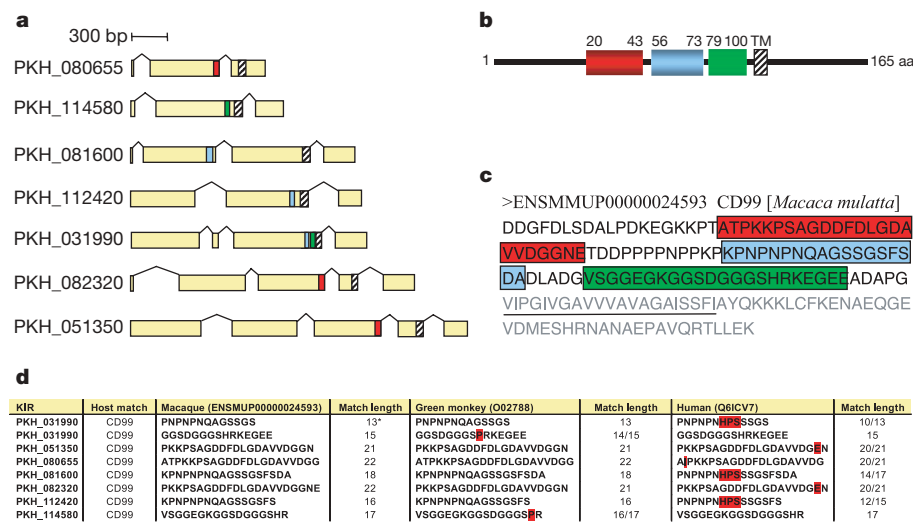


Figure 4 | Matches to CD99 host sequences in *P. knowlesi* (H strain).
a, Seven KIRs show conserved matches to three different regions of CD99 (shown in red, blue and green). **b**, Schematic view of *Macaca mulatta* CD99, showing matches to different KIRs. The numbers represent the amino acid position. TM, transmembrane domain. The highlighted regions represent the summary of perfectly matched amino acid stretches in the CD99 extracellular domain to a subgroup of seven KIR proteins. **c**, Amino acid sequence of *Macaca mulatta* CD99, highlighting the summary of matches to

KIRs. Amino acids corresponding to the transmembrane domain are underlined. The light-grey amino acids represent the transmembrane domain and the intracellular part of CD99. **d**, Comparison of the matches to *Macaca fascicularis*, African green monkey and human. Mismatches are highlighted in red. The asterisk refers to an additional host CD99 match in a KIR protein (PKH_031990) that did not satisfy the minimum length cutoff of 15 amino acids.

sequenced so far²⁷. In addition to the members of the PHIST family²⁷, an additional 100 proteins in *P. knowlesi* have typical PEXEL-like motifs (Supplementary Table 8 and Supplementary Fig. 22). Like the PfEMP1 protein in *P. falciparum*, the SICAs and KIRs lack a signal peptide and a typical PEXEL-motif. We have identified a novel motif in the N-terminal region of SICA- α domains with a positionally conserved tryptophan residue surrounded by hydrophilic residues (Supplementary Fig. 22) that may be the export signal. Similarly, 75% of KIR proteins have a conserved Z-L-P-S motif (where Z denotes a hydrophilic residue) at the beginning of the KIR domain that may also facilitate export (Supplementary Fig. 22). In summary, approximately 280 predicted *P. knowlesi* proteins may be exported to the infected erythrocyte surface via the PEXEL-dependent or PEXEL-independent pathways. By comparison, the exportome of *P. vivax* is considerably larger than that of *P. knowlesi* and seems to be much bigger than previously thought²⁷. About 145 *P. vivax* proteins contain typical PEXEL motifs including the members of the PHIST family and a small subgroup of 12 VIRs. Genome sequencing of *P. knowlesi* and its comparison with other malaria genomes has highlighted several novel features of this emerging and potentially life-threatening human malaria parasite, and underscores the importance of full genome sequencing of new *Plasmodium* species. Major differences in both content and organization of its genome were revealed that involve the host-parasite interface, reinforcing the notion that malaria species have evolved specific mechanisms for enhancing their survival within their respective hosts. The *P. knowlesi* genome will also greatly enhance the utility of this human-infective species as a model for addressing questions pertinent to all *Plasmodium* species.

METHODS SUMMARY

The random shotgun approach was used to obtain roughly eightfold coverage of the whole nuclear genome sequence from the erythrocyte stage of the Pk1(A+) clone of the H strain of *P. knowlesi*. Sequence reads were assembled (as described in the Supplementary Information) and positional information from sequenced read pairs were used to resolve the orientation and position of the contigs. The assembled *P. knowlesi* contigs were iteratively ordered and oriented by alignment to *P. vivax* assembled sequences (described in ref. 4) and by manual checking. Automated predictions from the gene finding algorithms were manually reviewed by comparison to orthologues in other *Plasmodium* species. Artemis and Artemis

Comparison Tool (ACT) were used (as described previously²⁸) for annotation and curation and viewing the TBLASTX comparisons of regions with conserved synteny between *P. knowlesi*, *P. vivax* and *P. falciparum*. This also allowed us to curate gene models and identify local interruptions of synteny. Functional annotations were based on standard protocols as described previously⁶.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 January; accepted 30 July 2008.

1. Cox-Singh, J. et al. *Plasmodium knowlesi* malaria in humans is widely distributed and potentially life-threatening. *Clin. Infect. Dis.* **46**, 165–171 (2008).
2. White, N. J. *Plasmodium knowlesi*: the fifth human malaria parasite. *Clin. Infect. Dis.* **46**, 172–173 (2008).
3. Brown, K. N. & Brown, I. N. Immunity to malaria: antigenic variation in chronic infections of *Plasmodium knowlesi*. *Nature* **208**, 1286–1288 (1965).
4. Carlton, J. M. et al. Comparative genomics of the neglected human parasite *Plasmodium vivax*. *Nature* doi:10.1038/nature07327 (this issue).
5. Howard, R. J., Barnwell, J. W. & Kao, V. Antigenic variation of *Plasmodium knowlesi* malaria: identification of the variant antigen on infected erythrocytes. *Proc. Natl Acad. Sci. USA* **80**, 4129–4133 (1983).
6. Gardner, M. J. et al. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**, 498–511 (2002).
7. Carlton, J. M. et al. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* **419**, 512–519 (2002).
8. Hall, N. et al. A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science* **307**, 82–86 (2005).
9. Janssen, C. S., Phillips, R. S., Turner, C. M. & Barrett, M. P. *Plasmodium* interspersed repeats: the major multigene superfamily of malaria parasites. *Nucleic Acids Res.* **32**, 5712–5720 (2004).
10. Berriman, M. et al. The genome of the African trypanosome *Trypanosoma brucei*. *Science* **309**, 416–422 (2005).
11. Barry, J. D. et al. What the genome sequence is revealing about trypanosome antigenic variation. *Biochem. Soc. Trans.* **33**, 986–989 (2005).
12. Nergadze, S. G., Rocchi, M., Azzalin, C. M., Mondello, C. & Giulotto, E. Insertion of telomeric repeats at intrachromosomal break sites during primate evolution. *Genome Res.* **14**, 1704–1710 (2004).
13. al-Khedery, B., Barnwell, J. W. & Galinski, M. R. Antigenic variation in malaria: a 3' genomic alteration associated with the expression of a *P. knowlesi* variant antigen. *Mol. Cell* **3**, 131–141 (1999).
14. Hiller, N. L. et al. A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science* **306**, 1934–1937 (2004).
15. Marti, M., Good, R. T., Rug, M., Knuepfer, E. & Cowman, A. F. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**, 1930–1933 (2004).
16. Finn, R. D. et al. The Pfam protein families database. *Nucleic Acids Res.* **36** (Database issue) D281–D288 (2008).

17. Barnwell, J. W., Howard, R. J., Coon, H. G. & Miller, L. H. Splenic requirement for antigenic variation and expression of the variant antigen on the erythrocyte membrane in cloned *Plasmodium knowlesi* malaria. *Infect. Immun.* **40**, 985–994 (1983).
18. Kajava, A. V. Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **277**, 519–527 (1998).
19. Galinski, M. R. & Corredor, V. Variant antigen expression in malaria infections: posttranscriptional gene silencing, virulence and severe pathology. *Mol. Biochem. Parasitol.* **134**, 17–25 (2004).
20. Deitsch, K. W., Calderwood, M. S. & Wellems, T. E. Malaria: Cooperative silencing elements in *var* genes. *Nature* **412**, 875–876 (2001).
21. Finlay, B. B. & McFadden, G. Anti-immunology: evasion of the host immune system by bacterial and viral pathogens. *Cell* **124**, 767–782 (2006).
22. Werner, E. B., Taylor, W. R. & Holder, A. A. A *Plasmodium chabaudi* protein contains a repetitive region with a predicted spectrin-like structure. *Mol. Biochem. Parasitol.* **94**, 185–196 (1998).
23. Goundis, D. & Reid, K. B. Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs. *Nature* **335**, 82–85 (1988).
24. Hall, R. *et al.* Mimicry of elastin repetitive motifs by *Theileria annulata* sporozoite surface antigen. *Mol. Biochem. Parasitol.* **53**, 105–112 (1992).
25. MacDonald, S. M. *et al.* Immune mimicry in malaria: *Plasmodium falciparum* secretes a functional histamine-releasing factor homolog *in vitro* and *in vivo*. *Proc. Natl Acad. Sci. USA* **98**, 10829–10832 (2001).
26. Urban, B. C. *et al.* *Plasmodium falciparum*-infected erythrocytes modulate the maturation of dendritic cells. *Nature* **400**, 73–77 (1999).
27. Sargeant, T. J. *et al.* Lineage-specific expansion of proteins exported to erythrocytes in malaria parasites. *Genome Biol.* **7**, R12 (2006).
28. Berriman, M. & Harris, M. Annotation of parasite genomes. *Methods Mol. Biol.* **270**, 17–44 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge the support of the Wellcome Trust Sanger Institute core sequencing and informatics groups. The study was funded by the Wellcome Trust through its support to the Pathogen Sequencing Unit at the Wellcome Trust Sanger Institute. We thank J. Barnwell for providing the Pk1(A+) clone of the H strain of the parasite for the generation of genomic DNA by A. Thomas. We thank A. Voorberg-vd Wel (BPRC, Rijswijk) for technical assistance. We thank D. Fergusson for providing us with the electron micrograph image of the erythrocyte, used in Fig. 2. Part of this work was supported by the Netherlands Organization for Scientific Research, NIH, BioMalPar and the Virimal contract. This work is dedicated to the memory of Marie-Adele Rajandream.

Author Contributions B.G.B., C.I.N., N.H., A.W.T. and C.M.R.T. initiated the project. M.A.Q., T.C., H.H., S.M., D.O., S.S., N.L., F.S., K.Br., R.S., S.T., S.M., M.Sa., M.Si., B.W. and D.W. constructed DNA libraries and performed sequencing; B.W., M.S. and I.C. finished and assembled sequence data; K.M., D. Harris and C.Ch. managed finishing and sequencing teams; M.A.R. managed the computational and bioinformatics support team; M.A.A., S.B., T.J.C., D. Harper, T.K., A.R.T., E.Z. and N.P. provided computational and bioinformatic support; U.B., A.E.B., E.M.P., S.L. and B.G.B. annotated the genome data. U.B., A.E.B., I.M.M., C.Ca., C.I.N., R.D.F., J.M., T.M., C.M.R.T., T.G.C., K.Bo., M.R.G., C.S.J., T.J.S., M.M., A.F.C., A.P.J., C.H.M.K., M.B. and A.P. contributed specific analysis topics presented in this manuscript or contributed data to characterize the genome and commented on manuscript drafts. U.B. performed data submission in EMBL. A.P., M.B., A.E.B., U.B. and C.I.N. drafted and edited the paper. A.P. and M.B. directed the project and A.P. assembled the manuscript.

Author Information The annotation and sequence data for the 14 chromosomes of the H strain of *P. knowlesi* have been submitted to the EMBL database with the following accession numbers: AM910983–AM910996. The annotation and sequence data are also available at <http://www.genedb.org> and <http://www.plasmodb.org>. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.P. (ap2@sanger.ac.uk).

METHODS

Parasite material and isolation of genomic DNA. Genomic DNA was isolated from blood drawn from an infected rhesus monkey at 10% ring stage parasitaemia. Blood was Plasmodipur-filtered five times to remove white blood cells and erythrocytes were lysed in 0.1% saponin. Total parasite DNA was isolated using the PUREGENE DNA isolation kit (Gentra Systems), according to the manufacturer's instructions. All experimental animal work in these studies was carried out under protocols approved by the independent Institutional Animal Care and Use Committee and performed according to Dutch and European laws.

Sequencing. We sequenced the *P. knowlesi* genome from plasmid clones containing small fragments of up to 4 kb inserted into pUC19 vector. Problems associated with high G+C sequence were addressed by optimizing the sequence mixture. The quality of reads for the project was as follows: 97.6% of *P. knowlesi* reads had a quality score of (derived from the PHRED score generated by GAP4²⁹) >70 ($P = 1 \times 10^{-7}$). Regions containing repeat sequences or an unexpected read depth were manually inspected. In addition, a *P. knowlesi* fosmid library was constructed in pCC1FOS vector and end sequences were produced (10.5-fold clone coverage) to obtain paired-end information from 40-kb inserts. In particular, we re-examined regions with apparent breaks in synteny for potential misassembly errors and location of several intrachromosomal telomeric-repeat (GGGTT[T/C]A) sequences associated with *SICAvar* and *kir* genes. Sequence reads were assembled with PHRED/PHRAP on the basis of overlapping sequence and were manually edited in GAP4 database²⁹. Information from oriented read pairs, together with additional sequencing from selected large-insert clones and synteny with *P. vivax* chromosomes, were used to resolve potential misassemblies. Using long-range sequence information from the fosmid end sequences, we were able to bridge 142 out of 190 total gaps (Supplementary Table 1).

Gene finding and genome annotation. Annotation (PK4 version of assembly) was performed using the Artemis³⁰ and ACT software³¹. Genes were identified by manual curation of the output of the gene finding software SNAP³² and Annotaid (an extension of the comparative gene prediction program Projector³³; I. M. Meyer, unpublished). A set of 100 manually curated *Plasmodium knowlesi* genes was used as the training set for SNAP predictions. Annotaid was optimized for genome-wide analysis by training its parameters with a manually curated training set of 180 orthologous gene pairs from *P. knowlesi* and *P. falciparum*.

Functional assignments were based on assessment of BLAST and FASTA similarity searches against public databases and searches in protein domain databases such as InterPro³⁴. In addition, TMHMMv2.0³⁵, SignalPv3.0³⁶ and t-RNA scan³⁷ were used to identify transmembrane domains, signal peptides and t-RNA genes.

To define the orthologous and paralogous relationships between the predicted proteomes of three *Plasmodium* species (*P. falciparum*, *P. knowlesi*, *P. vivax*), the OrthoMCL protein clustering algorithm³⁸ was used with an inflation value of 1.5.

To search for parasite proteins containing stretches of perfectly matched host sequences, the PMATCH algorithm (R. Durbin, unpublished) was used to report

exact matches of 15 amino acids or greater after screening out low complexity sequences (details are provided in Supplementary Information).

Building profile HMMs of SICA and KIR protein domains. Sequence alignments and dotter³⁹ analysis of SICA proteins revealed the presence of a distinct N-terminal cysteine-rich domain (termed SICA- α : in some cases there are two copies of this domain), multiple central cysteine-rich domains (SICA- β) and a C-terminal cytoplasmic encoding domain (SICA-C). For each domain, a profile HMM (using HMMer, <http://hmm.janelia.org/>) was constructed and searched against the *P. knowlesi* genome to find all examples of the domain (significant matches had *E*-values <0.001). The HMMs were rebuilt, using alignments constructed using all significant hits, and re-searched until no additional examples of the domain were found.

The program Phobius⁴⁰ was used to identify the putative transmembrane region located between the end of the last SICA- β domain and the SICA-C domain in all cases. An identical procedure was used to identify the domains in the KIR proteins. In this case, a single domain type was found on all KIR proteins, repeated between one and three times. Putative transmembrane proteins were identified as before, but only ~50% of KIR proteins had a predicted transmembrane region. Visual inspection of the corresponding C-terminal regions from sequences, both with and without predictions, showed the presence of a common hydrophobic patch. To investigate whether the Phobius⁴⁰ software was insufficiently sensitive to identify all of the KIR transmembrane regions, the predicted transmembrane regions were aligned and used to build a HMM of the transmembrane region. This was then used to iteratively search the whole genome as before.

29. Bonfield, J. K., Smith, K. & Staden, R. A new DNA sequence assembly program. *Nucleic Acids Res.* **23**, 4992–4999 (1995).
30. Rutherford, K. et al. Artemis: sequence visualization and annotation. *Bioinformatics* **16**, 944–945 (2000).
31. Carver, T. J. et al. ACT: the Artemis Comparison Tool. *Bioinformatics* **21**, 3422–3423 (2005).
32. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
33. Meyer, I. M. & Durbin, R. Gene structure conservation aids similarity based gene prediction. *Nucleic Acids Res.* **32**, 776–783 (2004).
34. Mulder, N. J. et al. InterPro, progress and status in 2005. *Nucleic Acids Res.* **33** (Database Issue) D201–D205 (2005).
35. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580 (2001).
36. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–795 (2004).
37. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
38. Li, L., Stoeckert, C. J. Jr & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).
39. Sonnhammer, E. L. & Durbin, R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**, GC1–GC10 (1995).
40. Kall, L., Krogh, A. & Sonnhammer, E. L. A combined transmembrane topology and signal peptide prediction method. *J. Mol. Biol.* **338**, 1027–1036 (2004).

LETTERS

Vancomycin-resistant enterococci exploit antibiotic-induced innate immune deficits

Katharina Brandl^{1†}, George Plitas², Coralia N. Mihu^{1†}, Carles Ubeda¹, Ting Jia¹, Martin Fleisher³, Bernd Schnabl^{4†}, Ronald P. DeMatteo² & Eric G. Pamer^{1,3}

Infection with antibiotic-resistant bacteria, such as vancomycin-resistant *Enterococcus* (VRE), is a dangerous and costly complication of broad-spectrum antibiotic therapy^{1,2}. How antibiotic-mediated elimination of commensal bacteria promotes infection by antibiotic-resistant bacteria is a fertile area for speculation with few defined mechanisms. Here we demonstrate that antibiotic treatment of mice notably downregulates intestinal expression of RegIII γ (also known as Reg3g), a secreted C-type lectin that kills Gram-positive bacteria, including VRE. Downregulation of RegIII γ markedly decreases *in vivo* killing of VRE in the intestine of antibiotic-treated mice. Stimulation of intestinal Toll-like receptor 4 by oral administration of lipopolysaccharide re-induces RegIII γ , thereby boosting innate immune resistance of antibiotic-treated mice against VRE. Compromised mucosal innate immune defence, as induced by broad-spectrum antibiotic therapy, can be corrected by selectively stimulating mucosal epithelial Toll-like receptors, providing a potential therapeutic approach to reduce colonization and infection by antibiotic-resistant microbes.

Infections caused by highly antibiotic-resistant bacteria, such as methicillin-resistant *Staphylococcus aureus* (MRSA) and VRE, are an increasing menace in hospitalized patients^{3,4}. Treatment of serious VRE infections is limited by the paucity of effective antibiotics⁵. VRE colonizes the gastrointestinal tract and it is likely that systemic bloodstream infections are the result of dissemination from the intestine⁶. It has been widely assumed that antibiotic treatment, by eliminating commensal flora, opens intestinal niches and provides increased access to nutrients, thereby enhancing VRE survival and proliferation. Recent studies, however, have demonstrated that commensal microbes in the intestine induce expression of proteins that restrict bacterial survival and growth^{7,8}. Thus, whereas commensal microbes may directly restrict VRE proliferation, an alternative hypothesis is that commensal microbes inhibit VRE indirectly by activating mucosal innate immune defenses.

To determine, as a first step, whether innate immune signalling by means of the Toll-like receptor (TLR)–MyD88 pathway restricts VRE colonization of the intestinal tract *in vivo*, we compared VRE killing in ligated ileal loops of wild-type and MyD88-deficient mice. Figure 1a demonstrates that VRE was killed much more effectively in ileal loops of wild-type than MyD88-deficient mice. Intestinal inhibition of VRE did not require MyD88-mediated signals in bone-marrow-derived cells, but instead was required in non-bone-marrow-derived, presumably epithelial, cells (Fig. 1b). RegIII γ is a secreted lectin with potent bactericidal activity against Gram-positive bacteria that is expressed by intestinal epithelial and Paneth cells⁷. Expression of RegIII γ is dependent on TLR–MyD88-mediated signals in intestinal epithelial cells and is induced by commensal microbes^{7,9}. To determine whether RegIII γ mediates *in vivo* killing

of VRE in the intestine, we injected a blocking polyclonal antiserum¹⁰ against RegIII γ into ileal loops of wild-type mice before inoculation of VRE. The number of surviving VRE bacteria was increased by over 400% in intestines treated with RegIII γ -specific antiserum (Fig. 1c), indicating that RegIII γ mediates *in vivo* VRE killing.

Administration of the broad-spectrum antibiotic combination metronidazole, neomycin and vancomycin (MNV), to which VRE is resistant, markedly increases colonization of the small intestine, caecum and colon with VRE (Fig. 2a and Supplementary Fig. 1), confirming previously published studies^{11–15}. It remains unclear, however, why elimination of the relatively sparse microbial flora of the small intestine increases the number of VRE by over 100–1,000-fold. Because VRE is acquired by the oral route and passes through the small intestine before systemic dissemination, we determined the impact of antibiotic administration on the expression of RegIII γ messenger RNA and protein levels in the distal ileum in wild-type mice and in mice treated with MNV for 1, 3 and 7 days. RegIII γ messenger RNA levels decreased within 1 day of antibiotic administration and continued to decrease over the subsequent 7 days,

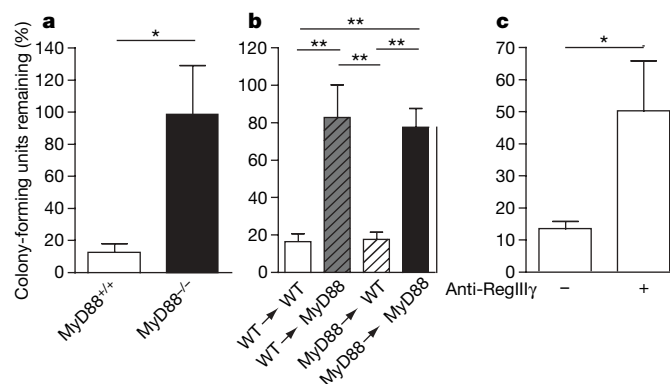


Figure 1 | MyD88-mediated signalling in non-haematopoietic cells is required for RegIII γ -mediated clearance of VRE. **a, b**, One-thousand VRE were injected into ileal loops of MyD88^{+/+}, MyD88^{-/-} (**a**) and MyD88 bone-marrow-chimaeric (**b**) mice. Two hours after injection the isolated section of the intestine was harvested and colony-forming units were determined; $n = 5–6$ each group. Bacterial numbers are expressed as the percentage of recovered VRE. For **a**, * $P = 0.02$, unpaired t -test; for **b**, ** $P < 0.01$, one-way ANOVA with Bonferroni correction. WT, wild type. **c**, Antiserum against RegIII γ or pre-immune serum was injected into ileal loops of wild-type mice before injection of 1,000 VRE. After 2 h, bacterial growth in the luminal fluid of the isolated section of the small intestine was determined. Bacterial numbers are expressed as the percentage of recovered VRE; $n = 5$ each group. * $P = 0.049$, unpaired t -test. Error bars denote s.e.m.

¹Infectious Diseases Service, Department of Medicine, Immunology Program, Sloan-Kettering Institute. ²Hepatobiliary Service, ³Department of Clinical Laboratories, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA. ⁴Department of Medicine, Columbia University, New York, New York 10032, USA. [†]Present addresses: Department of Genetics, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA (K.B.); Department of Infectious Diseases, MD Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA (C.N.M.); Department of Medicine, University of California San Diego, La Jolla, California 92093, USA (B.S.).

whereas RegIII γ protein expression was markedly diminished within 3 days of antibiotic initiation (Fig. 2b, c). Another antibiotic regimen consisting of enrofloxacin and clindamycin, which, like MNV, eliminates Gram-positive, Gram-negative and anaerobic commensal bacteria, similarly reduced RegIII γ mRNA and protein expression (Fig. 2d, e). In contrast, other mediators of antimicrobial defence—cryptdins and angiogenin-4—were not significantly downregulated by antibiotic administration (Supplementary Fig. 2).

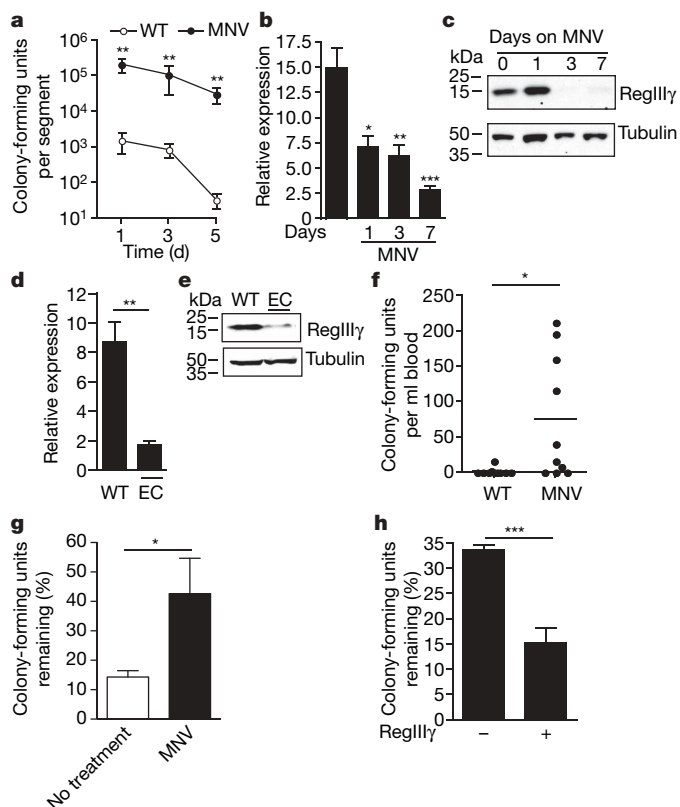


Figure 2 | RegIII γ expression is downregulated in antibiotic-treated mice, correlating with decreased clearance of VRE. **a**, Persistence and density of VRE intestinal colonization in mice. Mice were treated with MNV in drinking water starting 2 days before oral infection with 10^9 VRE by gavage. Bacterial counts within the distal small intestine were determined 1, 3 and 5 days after VRE infection ($n = 5$ each group and time point; $**P = 0.008$, Mann–Whitney test). **b**, Messenger RNA was extracted from the terminal ileum of wild-type mice (first bar) and mice receiving antibiotics (MNV) for 1, 3 and 7 days. RegIII γ expression was normalized to glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*; $n = 4$ –10 mice per group; $*P = 0.02$, $**P = 0.0011$, $***P = 0.0001$, each compared to untreated, unpaired *t*-test). **c**, Protein extracts from the distal small intestine of untreated and MNV-treated mice for 1, 3 and 7 days were analysed by western blotting with RegIII γ -specific antiserum. Tubulin was used as a loading control. **d**, Messenger RNA was extracted from the terminal ileum of untreated or enrofloxacin and clindamycin (EC)-treated mice, and RegIII γ expression was determined ($n = 2$ for wild-type and $n = 4$ for EC-treated mice; $**P = 0.0017$, unpaired *t*-test). **e**, Protein extracts from the distal small intestine of wild-type mice and mice receiving antibiotics (EC) for 7 days were analysed by western blotting with RegIII γ -specific antiserum. **f**, Wild-type mice and mice administered antibiotics (MNV) in the drinking water for 7 days were infected with 10^{10} VRE by gastric gavage, and the number of bacteria in the blood was determined 24 h after infection ($n = 10$ each group; $*P = 0.013$, Mann–Whitney test). **g**, Ileal loops of wild-type mice and mice receiving antibiotics for 7 days (MNV) were injected with 1,000 VRE, and bacterial survival was measured 2 h later. Bacterial numbers are expressed as percentage of recovered VRE ($n = 8$ each group; $*P = 0.04$, unpaired *t*-test). **h**, RegIII γ or control protein (β_2 -microglobulin) was inoculated into ileal loops, and VRE survival was assessed ($n = 5$ each group; $***P = 0.0004$, unpaired *t*-test). Error bars denote s.e.m.

To determine whether diminished RegIII γ levels correlate with increased susceptibility to VRE, wild-type mice receiving MNV were orally infected with VRE and dissemination to the blood stream was determined 24 h later. Whereas VRE bacteremia was uncommon in untreated mice, viable VRE was cultured from the blood of most antibiotic-treated mice (Fig. 2f). Moreover, assays for *in vivo* luminal killing demonstrated markedly impaired inactivation of VRE in the distal small bowel of antibiotic-treated mice (Fig. 2g). To determine whether reconstitution of RegIII γ in the intestine of antibiotic-treated mice corrects defective VRE killing, recombinant RegIII γ was injected into ileal loops of antibiotic-treated mice before VRE inoculation, and bacterial killing was measured. Administration of RegIII γ markedly diminished VRE survival in the gut of antibiotic-treated mice (Fig. 2h), demonstrating that exogenous RegIII γ complements the innate immune defect induced by broad-spectrum antibiotic therapy.

RegIII γ is induced in the mouse small intestine by the Gram-negative commensal bacterial species *Bacteroides thetaiotaomicron*¹⁶. In contrast, RegIII γ expression is downregulated by the Gram-positive, probiotic bacterium *Bifidobacterium longum*¹⁶. To determine whether TLR ligands associated with Gram-negative or Gram-positive bacteria can induce RegIII γ in antibiotic-treated mice, we added lipopolysaccharide (LPS) or lipoteichoic acid (LTA) to the drinking water of antibiotic-treated mice. Quantitative analysis of RegIII γ protein levels in distal ileum revealed significant upregulation after oral LPS treatment (Fig. 3b and Supplementary Fig. 3), and analysis of RegIII γ mRNA revealed a similar, albeit not statistically significant, trend (Fig. 3a). In contrast, oral administration of LTA to MNV-treated mice did not increase RegIII γ mRNA or protein (Fig. 3a, b and Supplementary Fig. 3).

To determine whether LPS-induced restoration of RegIII γ corrects the VRE clearance defect in the intestine of antibiotic-treated mice, luminal killing assays were performed in mice receiving MNV, in mice receiving MNV together with LPS or LTA, and in untreated mice. Whereas mice receiving only antibiotics manifested defective clearance of VRE from small intestine, bowel loops of mice receiving antibiotics together with LPS cleared VRE as efficiently as ileal loops of wild-type, untreated mice (Fig. 3c). Luminal killing of VRE in antibiotic-treated mice receiving LTA, in contrast, was not enhanced (Fig. 3c). The specificity and TLR4–MyD88-dependence of LPS-mediated rescue was investigated with MyD88- or TLR4-mutant mice. Antibiotic administration to MyD88-deficient mice did not increase VRE survival in the ileal lumen, and LPS administration did not increase VRE killing (Fig. 3c). LPS administration to TLR4-mutant mice also did not increase VRE killing in the intestinal lumen, demonstrating that the LPS effect in wild-type mice was TLR4-mediated and that LPS did not directly inhibit VRE survival (Fig. 3c and Supplementary Fig. 4). Antibiotic-treated mice had significantly higher numbers of VRE in the distal small intestine and mesenteric lymph nodes 24 h after oral infection compared to antibiotic-untreated mice, whereas no effect of LPS was found in TLR4-mutant mice (Supplementary Fig. 5a, b). LPS treatment reduced VRE colony-forming units in small intestines and mesenteric lymph nodes (MLNs) to levels observed in wild-type mice (Fig. 3d, e), demonstrating that LPS-mediated enhancement of innate immune defence decreased VRE colonization and dissemination. To exclude systemic LPS effects during oral LPS treatment, serum tumour necrosis factor (TNF) levels were determined in mice treated either with antibiotics or with antibiotics plus LPS for 7 days. TNF was not detected in either group, whereas mice receiving LPS intraperitoneally had increased serum TNF levels (Supplementary Fig. 6). To determine whether LPS-treated mice ingested equivalent amounts of antibiotic, vancomycin levels in the colons of mice receiving or not receiving LPS were determined. No significant difference was found between the groups, indicating that LPS administration did not reduce the intake of antibiotics in drinking water (Supplementary Fig. 7).

Whereas the preceding experiments demonstrated that administration of LPS concurrent with antibiotic treatment ameliorates the innate immune defect associated with elimination of the commensal

intestinal flora, in clinical situations it might be important to enhance resistance after antibiotics have already been administered. Therefore, mice that had received antibiotics for 4 days were treated with oral LPS. RegIII γ protein levels were significantly upregulated in mice receiving LPS after 4 days of antibiotic treatment and *RegIII γ* mRNA levels were also increased to levels that, however, did not achieve statistical significance (Fig. 4b and Supplementary Fig. 8). Immunohistochemistry demonstrated RegIII γ protein expression in Paneth cells and epithelial cells of the distal small intestine in wild-type

mice and an absence of RegIII γ in mice treated with antibiotics. In contrast, mice treated with LPS in addition to antibiotics expressed readily detectable amounts of RegIII γ on immunohistochemical analysis (Fig. 4c). In line with these data, luminal killing of VRE was also significantly enhanced in mice receiving LPS to rescue innate immune defences (Fig. 4d).

Our findings demonstrate that administration of broad-spectrum antibiotics compromises intestinal innate immune defences by eliminating commensal microbes, which diminishes expression of antimicrobial molecules in the intestinal mucosa. Compromised intestinal innate immune defence increases susceptibility to antibiotic-resistant microbes, including the Gram-positive bacterial pathogen VRE. Our experiments show that the innate immune effector molecule RegIII γ has an important role in host defence against VRE by killing bacteria in the small intestinal lumen. It is likely, however, that other MyD88-induced factors also inhibit intestinal colonization by VRE because MyD88-deficient mice were less effective at inhibiting VRE growth than wild-type mice treated with RegIII γ antiserum. Along similar lines, oral administration of LPS to antibiotic-treated mice, while inducing RegIII γ , may also increase intestinal resistance to VRE colonization by inducing other antimicrobial factors.

Altering the complexity of the intestinal commensal flora with broad-spectrum antibiotics provides an opportunity for otherwise innocuous microbes to cause systemic disease^{11,17,18}. Specific TLR signalling can substitute for signals induced by commensal microbes and restore RegIII γ , thereby enhancing defence against VRE. It is interesting that LPS, the principal component of the outer membrane of Gram-negative bacteria, induces RegIII γ , which acts selectively against Gram-positive bacteria by binding surface-exposed peptidoglycan. In contrast, LTA, a predominant surface glycolipid of Gram-positive bacteria, does not induce RegIII γ , nor does it enhance clearance of VRE. Our results, therefore, support the previous findings¹⁶ demonstrating that a commensal Gram-negative bacterium, but not a probiotic Gram-positive bacterium, induces RegIII γ , suggesting that bacteria

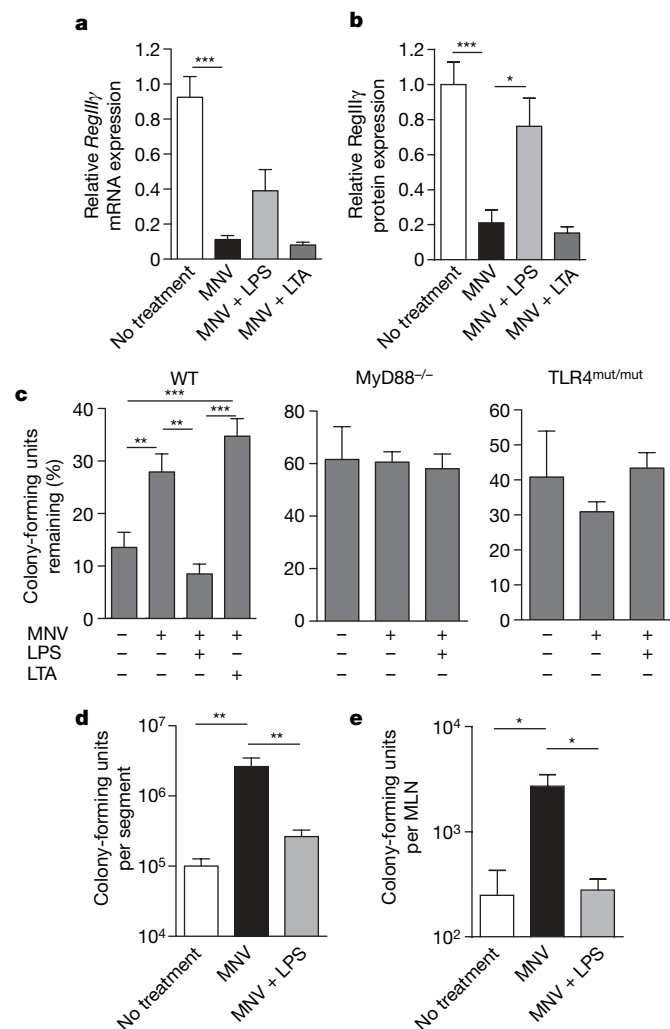


Figure 3 | Administration of LPS to antibiotic-treated mice restores RegIII γ levels and enhances luminal VRE killing. **a**, Mice were treated with MNV alone, MNV plus 2–4 $\mu\text{g l}^{-1}$ LPS, MNV plus 0.25 $\mu\text{g l}^{-1}$ LTA, or were left untreated, and *RegIII γ* mRNA expression was determined and is expressed relative to untreated mice ($n = 5$ –12 each group; *** $P < 0.001$, one-way ANOVA with Bonferroni correction). **b**, Quantitative western blot analysis with RegIII γ -specific antiserum was performed on protein extracts from the distal small intestine of untreated mice as well as mice receiving MNV, MNV plus LPS, or MNV plus LTA ($n = 5$ –12; * $P < 0.05$, *** $P < 0.001$, one-way ANOVA with Bonferroni correction). **c**, VRE survival was determined in ileal loops of wild type, MyD88 $^{-/-}$ or TLR4 $^{\text{mut/mut}}$ mice treated with MNV, MNV plus LPS, MNV plus LTA, or left untreated. Bacterial numbers are expressed as the percentage of recovered VRE ($n = 6$ –11 for wild type, $n = 7$ –8 for MyD88 $^{-/-}$, $n = 6$ –8 for TLR4 $^{\text{mut/mut}}$ (C3HeJ); ** $P < 0.01$, *** $P < 0.001$). **d**, Mice were treated for 7 days with MNV, MNV plus LPS, or left untreated, and then were orally infected with 10^{10} VRE. Twenty-four hours later, bacterial counts in the distal small intestine were determined ($n = 14$ –15; ** $P < 0.01$, one-way ANOVA with Bonferroni correction). **e**, Mice were treated as in **d** and, 24 h later, bacterial counts within MLNs were determined ($n = 4$ –5; * $P < 0.05$, one-way ANOVA with Bonferroni correction). Error bars denote s.e.m.

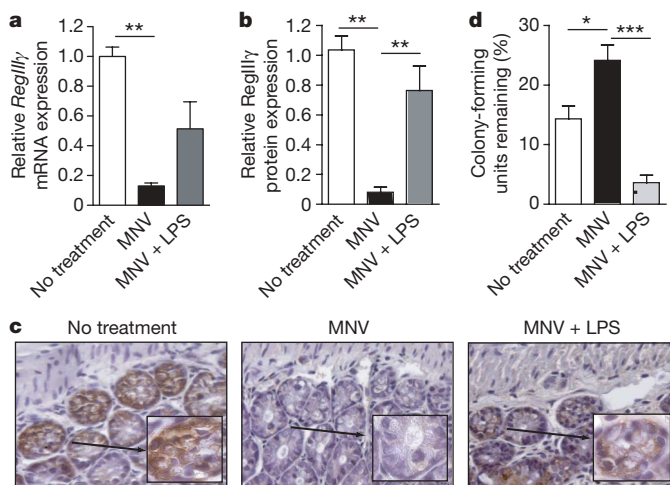


Figure 4 | Delayed LPS treatment of mice receiving antibiotics restores RegIII γ levels and enhances luminal VRE killing. **a**, Mice were treated with MNV for 4 days before the addition of LPS (1–4 $\mu\text{g l}^{-1}$) to the drinking water. *RegIII γ* mRNA induction was determined after 7 days LPS treatment and is expressed relative to untreated mice ($n = 9$ –15 in each group; ** $P < 0.01$, one-way ANOVA with Bonferroni correction). **b**, RegIII γ protein levels in the distal small intestine of mice treated as in **a** were determined ($n = 5$ –8; ** $P < 0.01$, one-way ANOVA with Bonferroni correction). **c**, Immunohistochemical detection of RegIII γ in paraffin-embedded distal small intestine sections in untreated mice, mice receiving MNV and mice receiving MNV plus LPS (magnification 400-fold, inset 1,000-fold). Representative samples from one of three mice is shown. **d**, VRE killing in ileal loops of mice treated as in **a** was determined and expressed as percentage of recovered VRE ($n = 5$ –6 each group; * $P < 0.05$, *** $P < 0.001$, one-way ANOVA with Bonferroni correction). Error bars denote s.e.m.

that belong to different classes can wage battles against each other by inducing innate antimicrobial mechanisms in the mammalian mucosa that selectively inactivate bacteria belonging to different classes.

Therapy with TLR ligands, though frequently contemplated, is complicated by the induction of inflammatory cytokines that can, at the extreme, result in septic shock and death¹⁹. This dire outcome occurs after systemic administration of some TLR ligands and thus profoundly limits enthusiasm for this form of therapy. In contrast, the gastrointestinal tract is exposed to TLR ligands with great regularity, if not continuously^{20,21}. Recent studies, however, indicate that the intestinal mucosa is not blind to TLR ligands^{22–24}, but that responses, while tempered, contribute to epithelial homeostasis^{25,26} and enhance innate immune defence against intestinal bacteria⁹. Thus, although systemic treatment with TLR ligands is fraught with difficulties, administration of TLR ligands to mucosal surfaces with the goal of enhancing innate immune defence may, in certain clinical situations, provide salutary effects.

Although it is presumed that depletion of commensal microbes has the direct effect of providing space and nutrients for pathogenic organisms, our findings indicate that broad-spectrum antibiotics may provide more important indirect advantages to pathogens by impairing mucosal innate immune defences. Specific induction of antimicrobial molecules, for example RegIII γ , by the oral administration of LPS could therefore be of potential therapeutic use in patients receiving broad-spectrum antibiotic therapy.

METHODS SUMMARY

Mice, bacteria and infection. Mice were infected by means of gavage with 10^9 or 10^{10} colony-forming units (as indicated) of the vancomycin-resistant *E. faecium* strain purchased from ATCC (stock number 700221). Bacterial counts within various organs were determined by homogenizing different organs in PBS containing 0.1% Triton X-100. Bacterial counts within luminal contents were determined by flushing the lumen with PBS containing 0.1% Triton X-100 and homogenizing the intestinal wall. For VRE counts in the blood, colony-forming units were determined by directly plating the blood. Plating was performed on Enterococcosel agar plates (Difco) with vancomycin ($8 \mu\text{g ml}^{-1}$, Sigma). VRE colonies were identified by appearance and confirmed by Gram staining.

Gut commensal depletion and reconstitution with LPS or LTA. Animals were provided with vancomycin (1 g l^{-1} , Sigma), neomycin sulphate (500 mg l^{-1} , Sigma) and metronidazole (1 g l^{-1} , Baxter) in drinking water for at least 7 days unless otherwise indicated. Alternatively, animals were given enrofloxacin (Baytril, 400 mg l^{-1}) and clindamycin (400 mg l^{-1} , Sigma) in the drinking water for 7 days. For some experiments, animals were treated for 4 days with vancomycin, neomycin and metronidazole before LPS from *Escherichia coli* 0127:B8 ($1\text{--}4 \mu\text{g } \mu\text{l}^{-1}$, Sigma; purified by phenol extraction) was added to the drinking water for an additional 7 days. Alternatively, for prophylactic administration, drinking water was supplemented from the beginning of the experiment on with vancomycin, neomycin and metronidazole together with LPS from *E. coli* 0127:B8 ($2\text{--}4 \mu\text{g } \mu\text{l}^{-1}$) or LTA from *Bacillus subtilis* ($0.25 \mu\text{g } \mu\text{l}^{-1}$, Sigma).

In vivo VRE killing. Ligated ileal loop experiments were performed as described previously⁹. PBS ($250 \mu\text{l}$) containing 1,000 VRE were injected into the ligated loop. Luminal fluid was plated after 2 h.

Generation of bone marrow chimaeras, RT-PCR, western blot, immunohistochemistry and determination of vancomycin levels were performed in accordance with standard protocols and as described previously⁹.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 May; accepted 14 July 2008.

Published online 24 August 2008.

1. Rice, L. B. Emergence of vancomycin-resistant enterococci. *Emerg. Infect. Dis.* **7**, 183–187 (2001).
2. Rice, L. B. Antimicrobial resistance in gram-positive bacteria. *Am. J. Infect. Control* **34**, S11–S19 (2006).
3. Murray, B. E. Vancomycin-resistant enterococcal infections. *N. Engl. J. Med.* **342**, 710–721 (2000).
4. Richards, M. J., Edwards, J. R., Culver, D. H. & Gaynes, R. P. Nosocomial infections in combined medical-surgical intensive care units in the United States. *Infect. Control Hosp. Epidemiol.* **21**, 510–515 (2000).

5. Wenzel, R. P. & Edmond, M. B. Managing antibiotic resistance. *N. Engl. J. Med.* **343**, 1961–1963 (2000).
6. Donskey, C. J. The role of the intestinal tract as a reservoir and source for transmission of nosocomial pathogens. *Clin. Infect. Dis.* **39**, 219–226 (2004).
7. Cash, H. L., Whitham, C. V., Behrendt, C. L. & Hooper, L. V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* **313**, 1126–1130 (2006).
8. Hooper, L. V., Stappenbeck, T. S., Hong, C. V. & Gordon, J. I. Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nature Immunol.* **4**, 269–273 (2003).
9. Brandl, K., Plitas, G., Schnabl, B., DeMatteo, R. P. & Pamer, E. G. MyD88-mediated signals induce the bactericidal lectin RegIII gamma and protect mice against intestinal *Listeria monocytogenes* infection. *J. Exp. Med.* **204**, 1891–1900 (2007).
10. Cash, H. L., Whitham, C. V. & Hooper, L. V. Refolding, purification, and characterization of human and murine RegIII proteins expressed in *Escherichia coli*. *Protein Expr. Purif.* **48**, 151–159 (2006).
11. Donskey, C. J. *et al.* Effect of antibiotic therapy on the density of vancomycin-resistant enterococci in the stool of colonized patients. *N. Engl. J. Med.* **343**, 1925–1932 (2000).
12. Miyazaki, S. *et al.* Development of systemic bacteraemia after oral inoculation of vancomycin-resistant enterococci in mice. *J. Med. Microbiol.* **50**, 695–701 (2001).
13. Stiefel, U., Pultz, N. J., Helfand, M. S. & Donskey, C. J. Increased susceptibility to vancomycin-resistant *Enterococcus* intestinal colonization persists after completion of anti-anaerobic antibiotic treatment in mice. *Infect. Control Hosp. Epidemiol.* **25**, 373–379 (2004).
14. Wells, C. L., Jechorek, R. P. & Erlandsen, S. L. Evidence for the translocation of *Enterococcus faecalis* across the mouse intestinal tract. *J. Infect. Dis.* **162**, 82–90 (1990).
15. Wells, C. L., Jechorek, R. P., Maddaus, M. A. & Simmons, R. L. Effects of clindamycin and metronidazole on the intestinal colonization and translocation of enterococci in mice. *Antimicrob. Agents Chemother.* **32**, 1769–1775 (1988).
16. Sonnenburg, J. L., Chen, C. T. & Gordon, J. I. Genomic and metabolic studies of the impact of probiotics on a model gut symbiont and host. *PLoS Biol.* **4**, e413 (2006).
17. Pultz, N. J., Stiefel, U., Subramanian, S., Helfand, M. S. & Donskey, C. J. Mechanisms by which anaerobic microbiota inhibit the establishment in mice of intestinal colonization by vancomycin-resistant *Enterococcus*. *J. Infect. Dis.* **191**, 949–956 (2005).
18. Kelly, C. P., Pothoulakis, C. & LaMont, J. T. *Clostridium difficile* colitis. *N. Engl. J. Med.* **330**, 257–262 (1994).
19. Kanzler, H., Barrat, F. J., Hessel, E. M. & Coffman, R. L. Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. *Nature Med.* **13**, 552–559 (2007).
20. Abreu, M. T., Fukata, M. & Arditi, M. TLR signaling in the gut in health and disease. *J. Immunol.* **174**, 4453–4460 (2005).
21. Lotz, M., Menard, S. & Hornef, M. Innate immune recognition on the intestinal mucosa. *Int. J. Med. Microbiol.* **297**, 379–392 (2007).
22. Cario, E. *et al.* Lipopolysaccharide activates distinct signaling pathways in intestinal epithelial cell lines expressing Toll-like receptors. *J. Immunol.* **164**, 966–972 (2000).
23. Hornef, M. W., Frisan, T., Vandewalle, A., Normark, S. & Richter-Dahlfors, A. Toll-like receptor 4 resides in the Golgi apparatus and colocalizes with internalized lipopolysaccharide in intestinal epithelial cells. *J. Exp. Med.* **195**, 559–570 (2002).
24. Ortega-Cava, C. F. *et al.* Strategic compartmentalization of Toll-like receptor 4 in the mouse gut. *J. Immunol.* **170**, 3977–3985 (2003).
25. Pull, S. L., Doherty, J. M., Mills, J. C., Gordon, J. I. & Stappenbeck, T. S. Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury. *Proc. Natl Acad. Sci. USA* **102**, 99–104 (2005).
26. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of commensal microflora by Toll-like receptors is required for intestinal homeostasis. *Cell* **118**, 229–241 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank L.V. Hooper for providing polyclonal RegIII antiserum, I. Leiner for technical assistance, and W. Falk, B. Salzberger and all members of the Pamer laboratory for discussions. This research was supported by the Alexander von Humboldt Foundation through a Feodor Lynen postdoctoral fellowship to KB and NIH grant AI39031 and AI42135 to EGP.

Author Contributions K.B., G.P., R.P.D. and E.G.P. designed the research. K.B., G.P., C.N.M., C.U., T.J., B.S. and M.F. performed the research. K.B. and E.G.P. analysed the data and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.G.P. (pamere@mskcc.org).

METHODS

Mice and bacteria. MyD88^{-/-} mice were provided by S. Akira. C57BL/6, C57BL/6 Ly5.1 and C3H/HeJ mice were purchased from Jackson Laboratory. All mice were kept and bred at Memorial Sloan-Kettering Cancer Center Research Animal Resource Center. All animal procedures were approved and performed according to institutional guidelines for animal care. All experiments were performed with the VRE from ATCC (stock number 700221), except for the experiment in Fig. 1a, which was done with a clinical VRE isolate from the microbiology laboratory at MSKCC.

Generation of bone marrow chimaeric mice. Recipient wild-type (CD45.1) or MyD88-deficient (CD45.2) mice were lethally irradiated with 950 cGy using a ¹³⁷Cs source and injected intravenously 2–3 h later with 5 × 10⁶ bone marrow cells derived from the tibias and femurs of the respective donors. Bone marrow chimaeric mice were used 7 weeks after engrafting.

Real-time PCR analysis. The small intestine was divided into four parts; a 1.5 cm segment was excised from the distal portion and total RNA was isolated using the Trizol reagent (Invitrogen). DNase-treated RNA underwent randomly primed cDNA synthesis and real-time PCR analysis. SYBR Green-based real-time PCR was performed using the DyNAmo SYBR Green qPCR Kit (Finnzymes). RegIIIγ- and angiogenin-4-specific primers were obtained from Qiagen. General cryptdin primer was used as described previously²⁷. Signals were normalized to *Gapdh* RNA (forward: 5'-ACCACAGTCCATGCCATCAC-3'; reverse: 5'-TCCACCACCTGTTGCTGTA-3'). Normalized data were used to quantify relative levels of RegIIIγ using ΔΔCt analysis.

Western blot analysis. Small intestines were divided into four pieces; samples were taken from the most distal part and western blotting using a polyclonal RegIIIγ antiserum or an anti-tubulin antibody (Santa Cruz) as loading control was performed as described previously¹⁰.

Western blot signals were analysed by densitometric measurements and subsequently quantified using the NIH-Image software program. Values obtained for RegIIIγ immunoblots were normalized to the optical density of corresponding immunoblots for tubulin.

Immunohistochemistry. Freshly isolated small intestine was divided into four parts. The fourth distal part was fixed in formalin and embedded in paraffin. Immunohistochemical staining for RegIIIγ was performed using the DAKO Envision System for rabbit primary antibodies according to the manufacturer's protocol. RegIIIγ polyclonal antiserum¹⁰ was diluted 1:500 in PBS containing 1% BSA. Control slides were stained with pre-immune serum or rabbit immunoglobulin G instead of the primary antibody and did not show any positive staining.

In vivo luminal killing. After anaesthesia a midline laparotomy incision was made. The intestine was occluded with a vascular clip 1 cm proximal to the ileo-caecal junction and 3 cm proximal to this point. Care was taken to avoid disrupting the mesenteric vascular arcades. The length of intestine between the two clips was injected with 250 μl PBS containing 1,000 VRE. After 2 h, mice were killed by exposure to carbon dioxide and the luminal fluid of the isolated segment was harvested and plated on Enterococcosc I agar plates. For RegIIIγ-blocking experiments, 250 μl of a 1:10 dilution of RegIIIγ polyclonal antiserum¹⁰ or pre-immune serum in PBS was injected into ileal loops of wild-type and antibiotic-treated mice. After 20 min, 250 μl PBS containing 1,000 VRE was injected. For reconstitution experiments, 250 μl MES-Buffer (25 mM, pH 6, 25 mM NaCl) containing 20 μM purified recombinant RegIIIγ was injected before adding 1,000 VRE in 100 μl MES-Buffer into the ileal loop. RegIIIγ was purified as described previously¹⁰. As a control, purified β2-microglobulin was substituted for RegIIIγ. After 2 h the animals were killed by exposure to carbon dioxide immediately before harvesting the isolated section of the small intestine.

Determination of the vancomycin levels. The colon was dissociated in 1 ml PBS and centrifuged for 10 min at 13,000g. Vancomycin levels were analytically determined by a particle-enhanced turbidimetric inhibition immunoassay method²⁸.

Array analysis and ELISA. Array analysis (Toll-like receptor signalling pathway PCR array from Superarray) was performed to determine the bioactivity of LPS and LTA in the specific concentrations used for the experiments (unpublished results). For intraperitoneal administration of LPS, mice were injected with 1 μg LPS. After 2 h, serum TNF levels were determined by ELISA (OptEIA kit from BD Pharmingen).

Statistical analysis. Statistical analysis was performed on Prism software. All *P*-values < 0.05 were considered to be significant. Error bars denote s.e.m.

27. Nenci, A. *et al.* Epithelial NEMO links innate immunity to chronic intestinal inflammation. *Nature* **446**, 557–561 (2007).

28. Hammett-Stabler, C. A. & Johns, T. Laboratory guidelines for monitoring of antimicrobial drugs. *Clin. Chem.* **44**, 1129–1140 (1998).

LETTERS

ATP drives lamina propria T_H17 cell differentiation

Koji Atarashi^{1*}, Junichi Nishimura^{1*}, Tatsuichiro Shima², Yoshinori Umesaki², Masahiro Yamamoto^{1,3}, Masaharu Onoue², Hideo Yagita⁴, Naoto Ishii⁵, Richard Evans⁶, Kenya Honda^{1,3} & Kiyoshi Takeda^{1,3}

Interleukin (IL)-17-producing CD4⁺ T lymphocytes (T_H17 cells) constitute a subset of T-helper cells involved in host defence and several immune disorders^{1,2}. An intriguing feature of T_H17 cells is their selective and constitutive presence in the intestinal lamina propria³. Here we show that adenosine 5'-triphosphate (ATP) that can be derived from commensal bacteria activates a unique subset of lamina propria cells, CD70^{high}CD11c^{low} cells, leading to the differentiation of T_H17 cells. Germ-free mice exhibit much lower concentrations of luminal ATP, accompanied by fewer lamina propria T_H17 cells, compared to specific-pathogen-free mice. Systemic or rectal administration of ATP into these germ-free mice results in a marked increase in the number of lamina propria T_H17 cells. A CD70^{high}CD11c^{low} subset of the lamina propria cells expresses T_H17-prone molecules, such as IL-6, IL-23p19 and transforming-growth-factor- β -activating integrin- α V and - β 8, in response to ATP stimulation, and preferentially induces T_H17 differentiation of co-cultured naive CD4⁺ T cells. The critical role of ATP is further underscored by the observation that administration of ATP exacerbates a T-cell-mediated colitis model with enhanced T_H17 differentiation. These observations highlight the importance of commensal bacteria and ATP for T_H17 differentiation in health and disease, and offer an explanation of why T_H17 cells specifically present in the intestinal lamina propria.

The intestinal mucosa has a unique and complicated immune system composed of a variety of cell populations. Among these, T_H17 cells, a subset of CD4⁺ T cells characterized by their STAT3-dependent expression of ROR γ t (encoded by *Rorc*) and production of IL-17, IL-22 and IL-21, control a variety of bacterial and fungal infections at mucosal surfaces^{1–7}. Importantly, aberrant T_H17 responses have been implicated in the pathogenesis of inflammatory bowel diseases⁸. The development of T_H17 cells has been shown to be controlled by the local cytokine milieu, including IL-6, transforming growth factor- β (TGF- β) and IL-23 (refs 1, 2, 7, 9–12). However, the mechanism of T_H17 development in the intestine is as yet not fully understood.

IL-17-expressing cells constitute a considerable proportion of CD4⁺ cells in the intestinal lamina propria, even in healthy mice kept under specific-pathogen-free (SPF) conditions (Supplementary Fig. 1a and ref. 3). The colonic lamina propria CD4⁺ cells also express messenger RNAs for IL-17, IL-17F and ROR γ t, representing the hallmarks of T_H17 cells (Supplementary Fig. 1b). The number of these 'naturally occurring' T_H17 cells in the colonic lamina propria increases with age (Supplementary Fig. 1c). Although interferon (IFN)- γ -positive CD4⁺ cells are similarly observed in the lamina propria and spleen, IL-17-producing cells are rarely observed in the spleen, mesenteric lymph node (MLN) or Peyer's patches (Supplementary Fig. 1a and ref. 3). Furthermore, the lamina propria

IL-17-producing CD4⁺ cells were normally observed in Peyer's-patch- and colonic-patch-null mice¹³ (Supplementary Fig. 2). These observations suggest that a specific environment in the lamina propria supports the generation of T_H17 cells *in situ*.

To investigate whether intestinal commensal bacteria are responsible for the generation of lamina propria T_H17 cells, we evaluated the numbers of T_H17 cells in germ-free mice. Although the numbers of colonic lamina propria CD4⁺ cells were not significantly changed (Supplementary Fig. 3a), the numbers of IL-17-positive CD4⁺ cells were greatly reduced in the large intestines of germ-free mice compared to those in SPF mice (Fig. 1a, b and Supplementary Fig. 3b). Consistent with previous reports¹⁴, the germ-free mice also exhibited severe reductions in their faecal IgA levels (Supplementary Fig. 3c), demonstrating that commensal bacteria contribute to the provision of a particular environment for lamina propria T_H17 cells as well as IgA-producing cells. To examine the role of commensal bacteria further, we treated SPF mice with a combination of vancomycin and metronidazole by oral administration, and analysed lamina propria T_H17 cells. The vancomycin- and metronidazole-treated mice showed marked reductions in both their faecal IgA levels and their numbers of IL-17-producing CD4⁺ cells (Supplementary Fig. 4a–c).

To assess the molecular basis for the commensal-bacteria-driven T_H17 differentiation, we examined the contribution of Toll-like receptor (TLR) signalling using *Myd88*^{−/−} *Trif*^{−/−} mice, which lack all TLR signalling. There was no detectable difference in the numbers of lamina propria IL-17-producing CD4⁺ cells between control and mutant animals (Fig. 1c, d and Supplementary Fig. 3d), indicating that the development of lamina propria T_H17 cells is independent of TLR signalling. It is worth noting that *Myd88*^{−/−} *Trif*^{−/−} mice showed impaired secretion of IgA in their faecal pellets (Supplementary Fig. 3e), indicating that the development of intestinal T_H17 cells and IgA-producing cells are both dependent on microflora, but are regulated by different mechanisms.

ATP has recently been shown to modulate immune cell functions by means of activation of the ATP sensors, P2X and P2Y receptors^{15–18}. In addition, bacteria have been shown to generate and secrete large amounts of ATP¹⁹. Indeed, ATP concentrations in faecal samples without bacterial lysis were much higher in SPF mice than in germ-free mice (Fig. 1e). Consistent with this result, ATP concentrations were reduced in faecal samples from SPF mice treated with vancomycin and metronidazole (Supplementary Fig. 4d). Furthermore, high ATP concentration was detected in the supernatant of *in vitro* cultured intestinal commensal bacteria derived from faeces of SPF mice (Supplementary Fig. 5). Therefore, although there might be other cellular sources of ATP such as dead epithelial cells, commensal bacteria may be a major source of intestinal luminal ATP. Interestingly, the faecal ATP concentrations were not reduced in

¹Laboratory of Immune Regulation, Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. ²Yakult Central Institute for Microbiological Research, Yaho 1796, Kunitachi, Tokyo 186-8650, Japan. ³WPI Immunology Frontier Research Center, Osaka University, Osaka 565-0871, Japan. ⁴Department of Immunology, Juntendo University School of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo 113-8421, Japan. ⁵Department of Microbiology and Immunology, Tohoku University Graduate School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai 980-8575, Japan. ⁶Department of Cell Physiology and Pharmacology, Henry Wellcome Building 2/59b, University of Leicester, Leicester LE1 9HN, UK.

*These authors contributed equally to this work.

Myd88^{-/-} *Trif*^{-/-} mice (Fig. 1f). These results prompted us to examine the contribution of ATP to the generation of intestinal T_H17 cells. To this end, we treated germ-free mice with a non-hydrolysable ATP analogue, ATP γ S (ref. 15), by intraperitoneal or rectal administration, and analysed the numbers of lamina propria IL-17-producing cells. The numbers of IL-17-producing CD4⁺ cells were markedly increased in the ATP γ S-treated germ-free mice (Fig. 1g–i). In contrast, ATP injection affected neither the numbers of IFN- γ -producing CD4⁺ cells (Fig. 1g, h) nor faecal IgA levels (Supplementary Fig. 3f). To assess further the possible involvement of ATP in T_H17 differentiation, we treated SPF mice with an ATP-hydrolysing enzyme, apyrase, or with an antagonist of P2X receptors, 2',3'-O-(2,4,6-trinitrophenyl)-ATP (TNP-ATP)¹⁵. In the apyrase- or TNP-ATP-treated mice, the numbers of IL-17-producing lamina propria CD4⁺ cells were significantly reduced (Fig. 1j, k), suggesting the key role of ATP in the generation of intestinal T_H17 cells.

Accumulating evidence suggests that lamina propria CD11c⁺ antigen-presenting cells directly sample the luminal contents and activate T cells^{20,21}. Indeed, the induction of mRNAs for IL-6 (encoded by *Il6*), IL-23p19 (*Il23a*) and integrin- α V (*Itgav*) and - β 8 (*Itgb8*) was observed in lamina propria CD11c⁺ cells, but not in CD11c⁻ cells or

epithelial cells, harvested from ATP γ S-treated germ-free mice (Supplementary Fig. 6). Integrin- α V and - β 8 are known to be involved in the activation of latent complexes of TGF- β ²². Therefore, we assumed that ATP promotes T_H17 differentiation by means of stimulation of lamina propria CD11c⁺ cells. Accordingly, we next co-cultured lamina propria CD11c⁺ cells with splenic naive CD4⁺CD62L⁺ T cells in the presence of a culture supernatant of intestinal commensal bacteria, which contained a high amount of ATP (Supplementary Fig. 5). Addition of this supernatant markedly enhanced differentiation of IL-17-expressing cells, but not of IFN- γ -expressing cells (Fig. 2a). This T_H17 differentiation was severely inhibited by the presence of apyrase. In contrast, a culture supernatant from *Salmonella typhimurium* showed a lower concentration of ATP and weaker ability to induce T_H17 differentiation (Fig. 2a and Supplementary Fig. 5). Notably, the culture supernatant from *S. typhimurium* potentially induced T_H1 differentiation, but this was not inhibited by apyrase (Fig. 2a).

To assess whether ATP was sufficient to induce T_H17 differentiation, we co-cultured lamina propria or splenic CD11c⁺ cells with splenic naive CD4⁺ cells in the presence of ATP γ S. The expression levels of *Il17a*, *Il17f*, *Il21* and *Il22*, but not *Ifng*, were markedly increased in CD4⁺ cells cultured with lamina propria CD11c⁺ cells in the presence of ATP γ S (Fig. 2b and Supplementary Fig. 7a). The T_H17 differentiation was also enhanced by another non-hydrolysable form of ATP, α , β -methylene-ATP (α β -ATP), and weakly enhanced by 2,3-O-(4-benzoylbenzoyl)-ATP (Bz-ATP)¹⁵ (Supplementary Fig. 7b). *Myd88*^{-/-} *Trif*^{-/-} CD11c⁺ cells and control cells exerted a similar effect on CD4⁺ cells, ruling out the possibility of endotoxin contamination (Supplementary Fig. 7c). The effect by ATP γ S or α β -ATP on T_H17 differentiation was inhibited by pharmacological blockade of P2X and P2Y receptors using suramin, or a combination of TNP-ATP and brilliant blue G¹⁵ (Supplementary Fig. 8). Importantly, these inhibitors had no effect

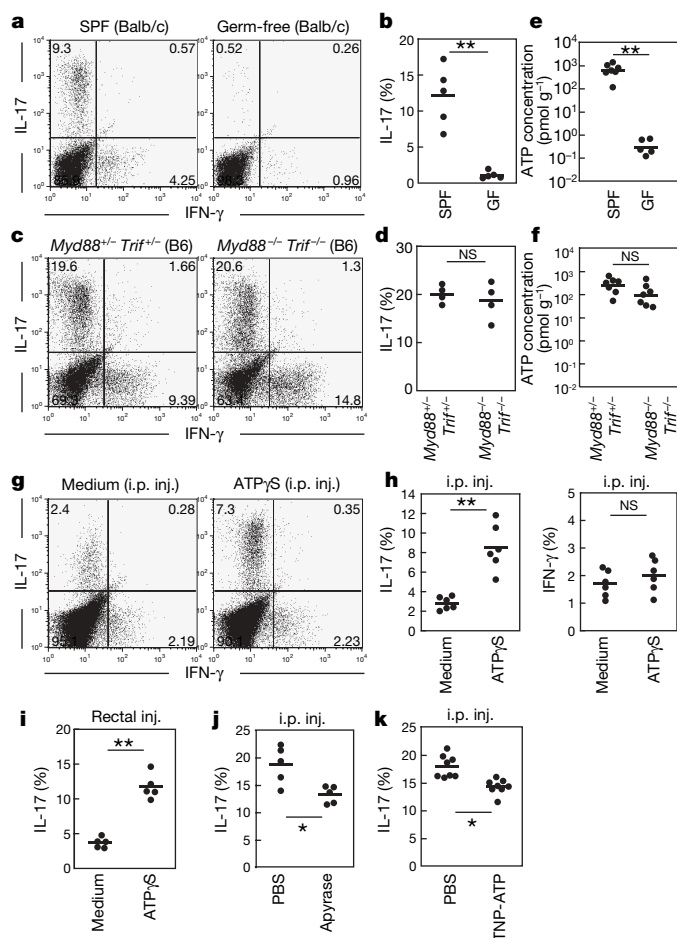


Figure 1 | Administration of ATP leads to a marked increase in lamina propria T_H17 cells in otherwise T_H17-lacking germ-free mice. **a–d**, Representative FACS dot plots gated on colonic lamina propria CD4⁺ cells in the indicated mice are shown in **a** and **c**, and the percentages of IL-17-producing CD4⁺ cells of individual mice are shown in **b** and **d**. GF, germ free. **e, f**, Faecal ATP levels (pmol per g faeces) in the indicated individual mice. **g–k**, Germ-free mice (ICR) were daily injected intraperitoneally (i.p.) or rectally with medium or ATP γ S (**g–i**). SPF mice were i.p. injected with PBS or apyrase, or with TNP-ATP (**j, k**). All mice were processed for FACS as in **a–d**. All experiments were performed more than twice with similar results. Horizontal bars indicate the means. ***P* < 0.01, **P* < 0.05; NS, not significant.

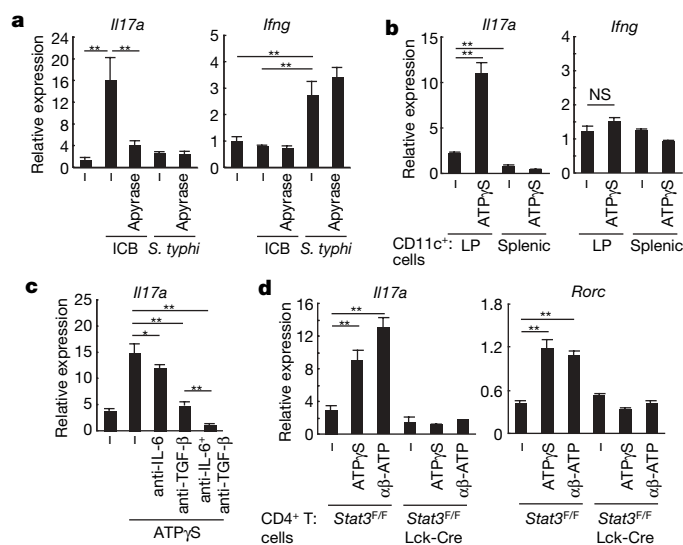


Figure 2 | ATP induces differentiation of naive CD4⁺ T cells into T_H17 cells. **a**, Splenic naive CD4⁺ T cells were co-cultured with colonic lamina propria CD11c⁺ with 20% conditioned medium from cultures of intestinal commensal bacteria (ICB) or *S. typhimurium* (*S. typhi*) in the presence or absence of apyrase. After 4 days, T cells were collected, restimulated and assayed for expression of *Il17* and *Ifng* by rRT-PCR. **b, c**, rRT-PCR analyses for the indicated genes of splenic naive CD4⁺ T cells cultured with colonic lamina propria (LP) or splenic CD11c⁺ cells in the presence of ATP γ S (**b**) with or without anti-IL-6, anti-TGF- β or their combination (**c**). **d**, rRT-PCR analyses for the indicated genes of splenic naive CD4⁺ T cells from *Stat3*^{F/F} Lck-Cre or *Stat3*^{F/F} mice cultured with wild-type colonic LP CD11c⁺ cells in the presence of ATP γ S or α β -ATP. Error bars, s.d. **P* < 0.05, ***P* < 0.01. Representative results are shown from one of two to five independent experiments.

on, or slightly enhanced, T_H1 cell differentiation. Although there might be other molecules affected by the ATP analogues and inhibitors we used, our results support the notion that the specific effect of ATP on T_H17 differentiation was mediated by P2X and P2Y receptors. The ATP-mediated enhancement of T_H17 differentiation was inhibited by anti-IL-6 or anti-TGF- β antibodies, and more severely inhibited by the combination of both antibodies (Fig. 2c). Consistent with this result, $Il6^{-/-}$ CD11c⁺ cells failed to induce the differentiation of IL-17-producing cells in response to ATP (Supplementary Fig. 7d). In addition, no effects of ATP on the differentiation of cells expressing *Il17a* and *Rorc* were observed for CD4⁺ cells from *Stat3^{E/F}* mice crossed onto Lck-Cre mice (*Stat3^{E/F}* Lck-Cre mice), in which the *Cre* transgene is under the control of the *Lck* promoter²³ (Fig. 2d). These results suggest that ATP stimulates lamina propria CD11c⁺

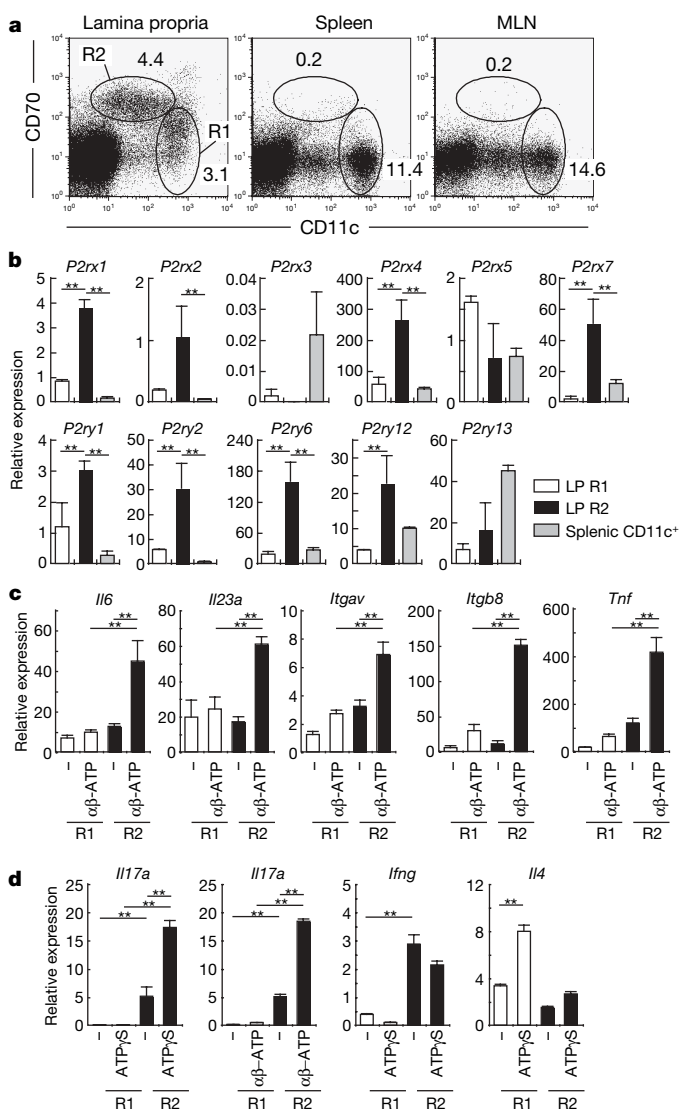


Figure 3 | A unique subset of lamina propria CD11c⁺ cells express P2X and P2Y receptors. **a**, Flow cytometry of cells isolated from the colonic lamina propria, spleen or MLN. Numbers indicate the percentages of CD70^{low}CD11c^{high} cells (R1) or CD70^{high}CD11c^{low} cells (R2). **b**, rRT-PCR analyses for P2X and P2Y receptors in R1, R2 or splenic CD11c⁺ cells. **c**, The sorted R1 and R2 cells were stimulated with $\alpha\beta$ -ATP for 3 h and assayed for expression of the indicated genes. **d**, The sorted R1 and R2 cells were co-cultured with splenic naive CD4⁺ T cells with ATP γ S or $\alpha\beta$ -ATP for 4 days. The levels of indicated mRNAs in the co-cultured T cells were analysed by rRT-PCR. Data are presented as means \pm s.d. of triplicate determinations. All experiments were performed more than three times with similar results. ** $P < 0.01$.

cells to produce IL-6 and TGF- β , thereby promoting the T_H17 differentiation by activation of STAT3 in CD4⁺ cells.

The heterogeneity of lamina propria CD11c⁺ cells has been addressed in a number of studies²⁴. Therefore, we investigated which subsets of lamina propria CD11c⁺ cells mediate T_H17 development. Flow cytometry analyses led to the identification of two major CD11c⁺ populations in the colonic lamina propria: CD70^{low}CD11c^{high} cells (termed R1) and CD70^{high}CD11c^{low} cells (termed R2; Fig. 3a). The R2 population comprised a unique subset present in the lamina propria (Fig. 3a). This subset was positive for myeloid lineage cell markers, such as F4/80 and CD11b, and exhibited immune-activating properties, as was evident from their high expression of CD70 and CD80, but was negative for the immune-suppressive marker CD103 (Supplementary Fig. 9a and refs 25 and 26). R2 cells also expressed CX3CR1; thus, this population seems to be a part of the previously described CX3CR1⁺CD11b^{high}CD11c^{low} T_H17 -inducing subset²⁶ (Supplementary Fig. 9a–c). R2 cells were observed in the lamina propria of germ-free mice or Peyer's-patch- and colonic-patch-null mice (Supplementary Fig. 11a, b), suggesting that this subset of cells develop by means of microflora- and Peyer's-patch-/colonic-patch-independent mechanisms. Because treatment with a blocking antibody for CD70 or deficiency of CD27, a receptor for CD70 (ref. 27), did not influence the number of lamina propria T_H17 cells (data not shown), CD70 itself is dispensable for T_H17 differentiation.

Next, we sorted R1 or R2 cells (Supplementary Fig. 10) and examined their expression of P2X and P2Y receptors by real-time polymerase chain reaction with reverse transcription (rRT-PCR). R2 cells expressed higher levels of mRNA for P2X1, P2X2, P2X4, P2X7, P2Y1, P2Y2, P2Y6 and P2Y12 receptors than R1 cells or splenic CD11c⁺ cells (Fig. 3b). Furthermore, R2 cells expressed *Il6*, *Il23a*, *Itgav* and *Itgb8* in response to ATP (Fig. 3c). In addition, on ATP stimulation, R2 cells expressed tumour necrosis factor- α (TNF- α), which has been implicated in T_H17 responses in humans²⁸. R2 cells, but not R1 cells, efficiently induced the differentiation of co-cultured naive CD4⁺CD62L⁺ T cells into IL-17-expressing cells, and this effect was markedly enhanced by the presence of ATP (Fig. 3d). Although the R2 population also induced T cells to differentiate into IFN- γ -positive cells, this effect was not enhanced by ATP γ S (Fig. 3d). In this experimental setting, IL-4-expressing cells were also induced, but this effect was more remarkable in co-cultures with R1 cells, suggesting that R1 and R2 cells carry out distinct missions with each other. R2 cells present in germ-free mice expressed low levels of *Il6* and *Itgb8*, and showed lower ability to induce T_H17 differentiation of co-cultured naive CD4⁺ T cells (Supplementary Fig. 11c, d); however, by the addition of ATP γ S, R2 cells from germ-free mice efficiently enhanced the T_H17 differentiation (Supplementary Fig. 11d). These results all support the hypothesis that commensal-bacteria-derived ATP is responsible for the activation of R2 cells and subsequent development of T_H17 cells.

To extend our understanding of the role of ATP in the differentiation of T_H17 cells, we examined the effects of ATP on intestinal inflammation. We adoptively transferred naive wild-type CD4⁺ T cells into severe combined immunodeficient (SCID) mice²⁹, treated these mice with $\alpha\beta$ -ATP or medium, and monitored the severity of their colitis. Although both groups developed colitis, treatment with $\alpha\beta$ -ATP exacerbated the symptoms of the disease, including diarrhoea and weight loss (Fig. 4a, Supplementary Fig. 12a and data not shown). Extensive oedema was prominent in the colon of $\alpha\beta$ -ATP-treated SCID mice (Supplementary Fig. 12b). Histological analyses revealed that inflammatory cell infiltration, epithelial hyperplasia and loss of goblet cells were more evident in the colon of $\alpha\beta$ -ATP-treated SCID mice (Fig. 4b). Notably, in contrast to a lack of considerable effect on IFN- γ -positive cells, the number of IL-17-positive CD4⁺ cells was significantly increased in the $\alpha\beta$ -ATP-treated SCID mice (Fig. 4c, d and Supplementary Fig. 12c). Thus, ATP-induced deterioration of colitis was accompanied by an increase in the

number of T_H17 cells in this T-cell-mediated colitis model, indicating the possible involvement of ATP in the generation of 'pathogenic' T_H17 cells.

From our results, we propose the following scenario: commensal bacteria-derived ATP activates $CD70^{high}CD11c^{low}$ cells in the lamina propria to induce IL-6 and IL-23 production as well as TGF- β activation, thereby leading to local differentiation of T_H17 cells. Our findings further suggest that this mechanism commonly operates during the differentiation of both 'naturally occurring' and 'pathogenic' T_H17 cells, although additional factors certainly contribute differentially to each case³⁰. The physiological nature and role of these T_H17 cells will be an interesting issue to be addressed in the context of the maintenance of intestinal mucosal homeostasis^{10,12}. The actual involvement of P2 receptors and subsequent intracellular signalling mechanisms responsible for the generation of the local cytokine milieu inducing T_H17 cells are also interesting issues for future studies. Finally, elucidating the entire picture of the regulation of the

development of intestinal regulatory T cells, T_H17 cells and other types of cells by ATP and its metabolites (ADP and adenosine) will provide valuable information towards our understanding of the complex system of intestinal mucosal immunity as well as the establishment of innovative ATP-targeted approaches for treating patients with inflammatory bowel diseases.

METHODS SUMMARY

Mice, cell isolation and faecal ATP measurements. C57BL/6J mice, CB-17 SCID mice and ICR germ-free mice were purchased from CLEA Japan. Balb/c germ-free mice were maintained at the Yakult Central Institute. *Myd88*^{-/-}, *Trif*^{-/-} and *Stat3*^{Y/F} Lck-Cre mice backcrossed eight or more generations onto C57BL/6J were used. The details of the procedures for isolation of lamina propria lymphocytes and $CD11c^+$ cells are described in Methods. Faeces from individual mice were weighed and suspended in PBS. The levels of ATP in the supernatants were determined with a luciferin-luciferase assay using an ATP assay kit (TOYO Ink.).

Intracellular cytokine staining and *in vitro* T-cell differentiation. The details of the procedures for intracellular cytokine staining and *in vitro* differentiation of $CD4^+$ T cells are described in Methods.

rRT-PCR. The details of the procedures and primers used for rRT-PCR are described in Methods. For all panels, bars represent the ratio of gene to *Gapdh* expression as determined by the relative quantification method ($\Delta\Delta CT$) (mean + s.d. of triplicate determination).

T-cell-mediated colitis model. Naive $CD4^+CD62L^{high}$ splenic T cells from Balb/c mice were intraperitoneally transferred into SCID mice (3×10^5 cells per mouse). The mice were then injected with $\alpha\beta$ -ATP (1.5 mg per mouse) or medium alone every 3 days for 4 weeks. After 6 or 8 weeks, colons were analysed.

Statistical analysis. Differences between control and experimental groups were evaluated using Student's *t*-test.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 June; accepted 7 July 2008.

Published online 20 August 2008.

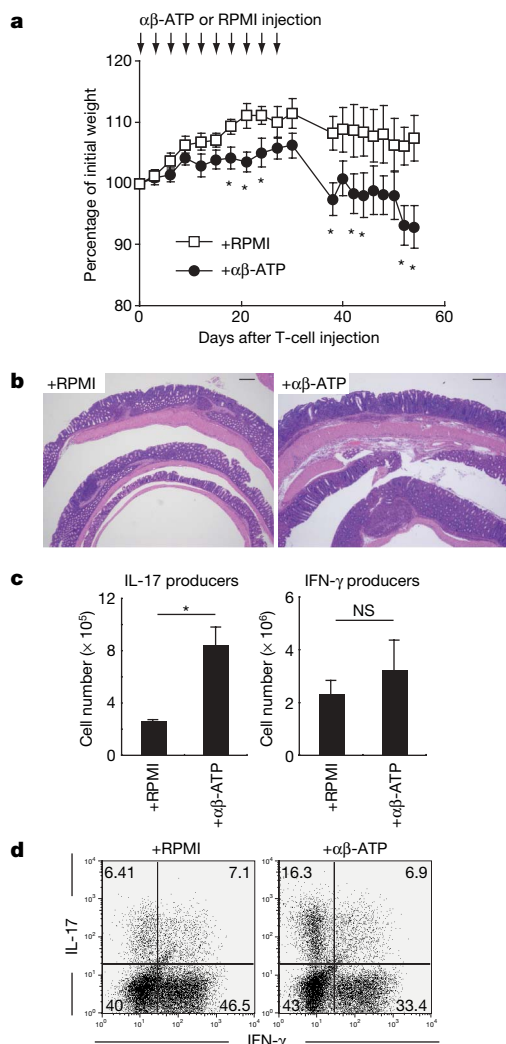


Figure 4 | Treatment with ATP exacerbates experimental colitis induced by adoptive transfer of naive $CD4^+$ T cells. SCID mice were transferred with naive $CD4^+$ T cells on day 0, and were intraperitoneally injected with $\alpha\beta$ -ATP (1.5 mg per mouse) or RPMI 1640 medium every 3 days for 1 month. **a**, The mean body weights of SCID mice injected with $\alpha\beta$ -ATP or RPMI ($n = 6$ per group) were monitored. Error bars, s.e.m. * $P < 0.05$ versus RPMI-injected mice. **b**, Haematoxylin- and eosin-stained colon sections of SCID mice treated with $\alpha\beta$ -ATP or RPMI. The numbers of IL-17- and IFN- γ -producing colonic lamina propria $CD4^+$ cells ($n = 3$ per group) are shown in **c**, and representative flow cytometry plots of $CD4^+$ T cells are shown in **d**. Error bars, s.d. * $P < 0.02$. NS, not significant. Two independent experiments were performed with similar results.

- Weaver, C. T., Hatton, R. D., Mangan, P. R. & Harrington, L. E. IL-17 family cytokines and the expanding diversity of effector T cell lineages. *Annu. Rev. Immunol.* **25**, 821–852 (2007).
- Bettelli, E., Oukka, M. & Kuchroo, V. K. T_H -17 cells in the circle of immunity and autoimmunity. *Nature Immunol.* **8**, 345–350 (2007).
- Ivanov, I. I. *et al.* The orphan nuclear receptor ROR γ t directs the differentiation program of proinflammatory IL-17⁺ T helper cells. *Cell* **126**, 1121–1133 (2006).
- Liang, S. C. *et al.* Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *J. Exp. Med.* **203**, 2271–2279 (2006).
- Nurieva, R. *et al.* Essential autocrine regulation by IL-21 in the generation of inflammatory T cells. *Nature* **448**, 480–483 (2007).
- Korn, T. *et al.* IL-21 initiates an alternative pathway to induce proinflammatory T_H 17 cells. *Nature* **448**, 484–487 (2007).
- Zhou, L. *et al.* IL-6 programs T_H -17 cell differentiation by promoting sequential engagement of the IL-21 and IL-23 pathways. *Nature Immunol.* **8**, 967–974 (2007).
- Cho, J. H. & Weaver, C. T. The genetics of inflammatory bowel disease. *Gastroenterology* **133**, 1327–1339 (2007).
- Veldhoen, M., Hocking, R. J., Atkins, C. J., Locksley, R. M. & Stockinger, B. TGF β in the context of an inflammatory cytokine milieu supports *de novo* differentiation of IL-17-producing T cells. *Immunity* **24**, 179–189 (2006).
- Mangan, P. R. *et al.* Transforming growth factor- β induces development of the T_H 17 lineage. *Nature* **441**, 231–234 (2006).
- Bettelli, E. *et al.* Reciprocal developmental pathways for the generation of pathogenic effector T_H 17 and regulatory T cells. *Nature* **441**, 235–238 (2006).
- Cua, D. J. & Kastelein, R. A. TGF- β , a 'double agent' in the immune pathology war. *Nature Immunol.* **7**, 557–559 (2006).
- Chang, S. Y. *et al.* Colonic patches direct the cross-talk between systemic compartments and large intestine independently of innate immunity. *J. Immunol.* **180**, 1609–1618 (2008).
- Fagarasan, S. & Honjo, T. Intestinal IgA synthesis: regulation of front-line body defences. *Nature Rev. Immunol.* **3**, 63–72 (2003).
- North, R. A. Molecular physiology of P2X receptors. *Physiol. Rev.* **82**, 1013–1067 (2002).
- Schnurr, M. *et al.* Extracellular nucleotide signaling by P2 receptors inhibits IL-12 and enhances IL-23 expression in human dendritic cells: a novel role for the cAMP pathway. *Blood* **105**, 1582–1589 (2005).
- Khakh, B. S. & North, R. A. P2X receptors as cell-surface ATP sensors in health and disease. *Nature* **442**, 527–532 (2006).
- Idzko, M. *et al.* Extracellular ATP triggers and maintains asthmatic airway inflammation by activating dendritic cells. *Nature Med.* **13**, 913–919 (2007).

19. Ivanova, E. P., Alexeeva, Y. V., Pham, D. K., Wright, J. P. & Nicolau, D. V. ATP level variations in heterotrophic bacteria during attachment on hydrophilic and hydrophobic surfaces. *Int. Microbiol.* **9**, 37–46 (2006).
20. Rescigno, M. *et al.* Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature Immunol.* **2**, 361–367 (2001).
21. Niess, J. H. *et al.* CX3CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance. *Science* **307**, 254–258 (2005).
22. Travis, M. A. *et al.* Loss of integrin $\alpha(v)\beta8$ on dendritic cells causes autoimmunity and colitis in mice. *Nature* **449**, 361–365 (2007).
23. Takeda, K. *et al.* Stat3 activation is responsible for IL-6-dependent T cell proliferation through preventing apoptosis: generation and characterization of T cell-specific Stat3-deficient mice. *J. Immunol.* **161**, 4652–4660 (1998).
24. Iwasaki, A. Mucosal dendritic cells. *Annu. Rev. Immunol.* **25**, 381–418 (2007).
25. Coombes, J. L. *et al.* A functionally specialized population of mucosal CD103⁺ DCs induces Foxp3⁺ regulatory T cells via a TGF- β and retinoic acid-dependent mechanism. *J. Exp. Med.* **204**, 1757–1764 (2007).
26. Denning, T. L., Wang, Y. C., Patel, S. R., Williams, I. R. & Pulendran, B. Lamina propria macrophages and dendritic cells differentially induce regulatory and interleukin 17-producing T cell responses. *Nature Immunol.* **8**, 1086–1094 (2007).
27. Hendriks, J. *et al.* CD27 is required for generation and long-term maintenance of T cell immunity. *Nature Immunol.* **1**, 433–440 (2000).
28. Zaba, L. C. *et al.* Amelioration of epidermal hyperplasia by TNF inhibition is associated with reduced Th17 responses. *J. Exp. Med.* **204**, 3183–3194 (2007).
29. Leach, M. W., Bean, A. G., Mauze, S., Coffman, R. L. & Powrie, F. Inflammatory bowel disease in C.B-17 scid mice reconstituted with the CD45RB^{high} subset of CD4⁺ T cells. *Am. J. Pathol.* **148**, 1503–1515 (1996).
30. Mazmanian, S. K., Round, J. L. & Kasper, D. L. A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature* **453**, 620–625 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank A. Iwasaki and N. Tsuji for discussion, M. H. Jang, H. Ohno, H. Yamane, M. Yoshida and H. Shiomi for technical advice and reagents, and J. Borst for CD27-deficient mice. This work was supported by Grants-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology, the Ministry of Health, Labour and Welfare, the Osaka Foundation for the Promotion of Clinical Immunology, the Ichiro Kanehara Foundation, Sumitomo Foundation, Senri Life Science Foundation and the Naito Foundation.

Author Contributions K.H. conceived the research, planned experiments and analyses and wrote the paper; K.A. and J.N. largely conducted experiments; T.S. performed some of the experiments; Y.U., M.Y., M.O., H.Y., N.I. and R.E. provided key materials; and K.T. oversaw the whole project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.H. (honda@ongene.med.osaka-u.ac.jp) or K.T. (ktakeda@ongene.med.osaka-u.ac.jp).

METHODS

Reagents. ATP γ S (A1388), $\alpha\beta$ -ATP (M6517) and Bz-ATP (B6396) were purchased from Sigma-Aldrich, and dissolved in phenol-red-free RPMI1640 medium. Apyrase from potato (ATPase activity: 40–200 units (U) per mg protein, A7646), suramine (S2671), TNP-ATP (T4193) and BBG (B0770) were purchased from Sigma-Aldrich, and dissolved in PBS. For *in vivo* experiments, 80 U apyrase or 250 μ g of TNP-ATP was dissolved in 300 μ l PBS and immediately injected intraperitoneally into SPF mice every 3 days for 15 days; or 1.25 mg of ATP γ S was dissolved in phenol-red-free RPMI1640 medium and injected daily intraperitoneally or intrarectally with syringes or 1,000- μ l pipette tips, respectively, into germ-free mice for 6 days. For *in vitro* experiments, ATP γ S, $\alpha\beta$ -ATP and Bz-ATP were used at 5–10 μ M; suramin at 30 μ M; TNP-ATP at 10 μ M; BBG at 100 nM; and apyrase at 15 U ml⁻¹. The anti-IL-6 and anti-TGF- β antibodies were purchased from R&D Systems and used at 10 μ g ml⁻¹ or 25 μ g ml⁻¹, respectively.

Mice. CB-17 SCID mice and ICR germ-free mice were purchased from CLEA Japan. Balb/c germ-free mice were maintained at the Yakult Central Institute. *Myd88*^{-/-}, *Trif*^{-/-} and *Stat3*^{EVF} Lck-Cre mice were generated as described previously^{23,31}, and backcrossed eight or more generations onto C57BL/6J. CD27-deficient mice were used with permission from J. Borst²⁷. Unless otherwise indicated, wild-type C57BL/6J mice purchased from CLEA Japan maintained under SPF conditions were used. All experiments were performed in accordance with the Guidelines for Animal Experiments of Osaka University.

Faecal ATP measurements. Faeces from individual mice were collected, weighed and gently suspended in PBS containing 0.01% NaN₃. After centrifugation, the supernatants were collected and passed through 0.22- μ m filters. The levels of ATP were determined with a luciferin-luciferase assay using an ATP assay kit (TOYO Ink) according to the manufacturer's instructions, except that the cell lysis step was omitted.

Bacterial culture. Faeces were collected from SPF mice, gently dissolved in 10 ml of serum-free RPMI medium, passed through a mesh and incubated under aerobic condition at 37 °C for 16 h. A small aliquot of the faecal suspension was transferred into 3 ml fresh medium and further incubated under aerobic conditions at 37 °C. The growth phases of these culturable intestinal aerobic commensal bacteria were monitored spectrophotometrically at 590 nm, and ATP concentrations in supernatants were measured as above. *S. typhimurium* (obtained from the Research Institute for Microbial Diseases bacterial culture collection, Osaka, Japan) was grown in the same medium, and ATP concentration was measured in the same way. The culture supernatants collected at the late exponential phase of bacterial growth were passed through 0.22- μ m filters and used as conditioned media for co-culture of CD11c⁺ cells and naive CD4⁺ T cells.

Isolation of lymphocytes. To prepare single-cell suspensions from spleens, mesenteric lymph nodes and Peyer's patches, the collected organs were ground between glass slides, and the cells were passed through 40- μ m nylon meshes and suspended in HBSS. Splenocytes were treated with RBC lysis buffer (0.15 M NH₄Cl, 1 mM KHCO₃, 0.1 mM EDTA) for 5 min before suspension. Naive CD4⁺ T cells were purified from spleens using a CD4⁺CD62L⁺ T cell isolation kit II (Miltenyi Biotec; purity 95%). For isolation of lamina propria lymphocytes (see also ref. 3), intestines were opened longitudinally, washed to remove faecal content, and shaken in HBSS containing 5 mM EDTA for 20 min at 37 °C. After removal of epithelial cells and fat tissue, the intestines were cut into small pieces and incubated with RPMI1640 containing 4% fetal bovine serum (FBS), 1 mg ml⁻¹ collagenase type II (Invitrogen), 1 mg ml⁻¹ dispase (Invitrogen) and 40 μ g ml⁻¹ DNase I (Roche Diagnostics) for 1 h at 37 °C in a shaking water bath. The digested tissues were washed with HBSS containing 5 mM EDTA, resuspended in 5 ml of 40% Percoll (GE Healthcare) and overlaid on 2.5 ml of 80% Percoll in a 15-ml Falcon tube. Percoll gradient separation was performed by centrifugation at 780g for 20 min at 25 °C. The lamina propria lymphocytes were collected at the interface of the Percoll gradient and washed with MACS buffer (0.5% BSA and 2 mM EDTA in PBS) or RPMI1640. The cells were used immediately for experiments.

Isolation of lamina propria CD11c⁺ cells. Lamina propria CD11c⁺ cells were isolated using a protocol modified from an EDTA perfusion method³². Mice were anesthetized and their peritoneal and pleural cavities opened. Next, 10 ml of 20 mM EDTA in HBSS was perfused through the left ventricle of the heart. At the end of the perfusion, the entire colon or small intestine excluding the caecum was removed, opened longitudinally and washed to remove faecal content. Colons were then cut into halves and placed in tubes filled with 2 mM EDTA in HBSS. The tubes were shaken at 4,800 oscillations per min for 50 s using a mini beater (Biospec Products). The tissues were washed with PBS and the epithelial cells and muscle layers removed with tweezers. In some experiments, the epithelial cell sheets were collected and kept in ice-cold HBSS until RNA extraction. After

removal of epithelial cells and muscle layers, the tissues were then cut into small pieces and incubated with RPMI1640 containing 4% FBS, 1 mg ml⁻¹ collagenase type II, 1 mg ml⁻¹ dispase and 40 μ g ml⁻¹ DNase I for 1 h at 37 °C in a shaking water bath. After filtration of the digested tissues through a 40- μ m cell strainer, the isolated cells were washed with MACS buffer and CD11c-positive cells were purified twice to >95% purity using CD11c MicroBeads (Miltenyi Biotec). In some experiments, enriched CD11c-positive cells were further stained with specific antibodies and sorted by FACS Vantage SE (BD Biosciences), with a resulting purity of around 95%.

Isolation of splenic CD11c⁺ cells. Spleens were collected and incubated with RPMI1640 containing 4% FBS, 1 mg ml⁻¹ collagenase type II, 1 mg ml⁻¹ dispase and 40 μ g ml⁻¹ DNase I for 1 h at 37 °C in a shaking water bath. After treatment with RBC lysis buffer, the isolated cells were washed with MACS buffer and CD11c-positive cells purified twice to >95% purity using CD11c MicroBeads (Miltenyi Biotec).

Intracellular cytokine staining. The intracellular expression of IL-17 and IFN- γ in CD4⁺ T cells was analysed using a Cytotfix/Cytoperm Kit Plus (with GolgiStop; BD Biosciences) according to the manufacturer's instructions. In brief, lymphocytes obtained from the intestinal lamina propria, spleens, mesenteric lymph nodes or Peyer's patches were incubated with 50 ng ml⁻¹ phorbol myristate acetate (PMA; Sigma), 5 μ M calcium ionophore A23187 (Sigma) and GolgiStop in complete media at 37 °C for 4 h. Surface staining was performed with a corresponding cocktail of fluorescently labelled antibodies for 20 min at 4 °C; after this, the cells were permeabilized with Cytotfix/Cytoperm solution for 20 min at 4 °C and intracellular cytokine staining was performed with fluorescently labelled cytokine antibodies for 20 min.

In vitro T-cell differentiation. Naive CD4⁺ T cells were collected for 4 days with purified CD11c⁺ cells and 0.5 μ g ml⁻¹ anti-CD3 antibody (BD Biosciences) in the presence or absence of ATP. The cultured cells were harvested and rested for 2 h before being restimulated with PMA and calcium ionophore (Sigma) for 3 h.

Flow cytometry. The following antibodies were used: anti-CD4-CyChrom, anti-IFN- γ -FITC, anti-IL-17-PE, anti-CD8-FITC, anti-CD11b-PE, anti-CD11b-FITC, anti-CD11c-PE, anti-CD11c-APC, anti-CD80-FITC, anti-CD103-FITC (all BD Biosciences), anti-CCR6-PE (R&D Systems Inc.), anti-F4/80-Alexa488 (Caltag Laboratories), anti-CD70-biotin (eBioscience), rabbit anti-CX3CR1 antibody (Torrey Pines Biolabs), streptavidin-Alexa488 and goat anti-rabbit IgG-Alexa488 (Molecular Probes). Data were acquired using a FACSCalibur or a FACSCanto (BD Biosciences) and analysed using Flowjo software (Tree Star).

Real-time RT-PCR. Complementary DNAs were synthesized from RNA samples prepared with an RNeasy Mini Kit (Qiagen) using Rever Tra Ace (Toyobo). Complementary DNAs were analysed by rRT-PCR using Power SYBR Green PCR Master Mix (Applied Biosystems) in an ABI 7300 real time PCR system (Applied Biosystems). Serial dilutions of a standard were included for each gene to generate a standard curve and allow calculation of the input amount of cDNA for each gene. Values were then normalized by the amount of GAPDH in each sample. The primer sets were designed with Primer Express Version 3.0 (Applied Biosystems) and initially tested to confirm comparable (>90%) efficiencies. The following primer sets were used: *Il17a*, 5'-GGACTCTCCACCGCAATGA-3' and 5'-GGCACTGAGCTTCCAGATC-3'; *Il17f*, 5'-CCCCATGGGATTTACAACATCAC-3' and 5'-CATTGATGCAGCCTGAGTGTCT-3'; *Il21*, 5'-GGCAATGAAAGCCTGTGGAA-3' and 5'-GGCAATGAAAGCCTGTGGAA-3'; *Rorc*, 5'-GGAGGACAGGGAGCCAAAGTT-3' and 5'-CCGTAGTGGATCCAGATGACT-3'; *Ifng*, 5'-TCAAGTGGCATAGATGTGGAAGAA-3' and 5'-TGGCTCTGCAGGATTTTCATG-3'; *Il4*, 5'-GGCATTTTGAACAGAGGT-CACA-3' and 5'-GACGTTTGGCACATCCATCTC-3'; *Il6*, 5'-CTGCAAGAGACTTCCATCCAGTT-3' and 5'-AAGTAGGGAAGGCCGTGGTT-3'; *Itgb8*, 5'-ACAGCATCGCATGGACCA-3' and 5'-AAGCAACCCGATCAA-GAATGTG-3'; *Itgav*, 5'-CGCCTATCTTCGGGATGAATC-3' and 5'-CCAACCGTAAGTCCATGAAATG-3'; *Gapdh*, 5'-CCTCGTCCCGTAGA-CAAAATG-3' and 5'-TCTCCACTTTGCCACTGCAA-3'; *P2rx1*, 5'-ACGAAACAAGAAGTGGGAGT-3' and 5'-AGGCCACTTGAGGTCTGGTAT-3'; *P2rx2*, 5'-GAGAGCTCCATCATCACCAAA-3' and 5'-CAGGGTCTGGGA-AGGAGTAAC-3'; *P2rx3*, 5'-CCGAGAAGCTTCACTTTTCA-3' and 5'-TTATGTCTTGTTCGGTGAGG-3'; *P2rx4*, 5'-TGCTGACAAATTCAGG-TTTC-3' and 5'-GATCATGGTTGGGATGATGTC-3'; *P2rx5*, 5'-AACCGTCTGGACAACAAACAC-3' and 5'-TTTCATCAGGTACGGAATC-3'; *P2rx7*, 5'-CCAGGAAGCAGGAGAGAACTT-3' and 5'-ATCCGTGTTCTT-GTCAATCCAG-3'; *P2ry1*, 5'-GGCAGGCTCAAGAAGAAGAT-3' and 5'-TCCCATGTCAGAGTAGAAGA-3'; *P2ry2*, 5'-CCGAGAGACTCTTTAGC-CATTT-3' and 5'-GCCATAAGCACGTAACAGACC-3'; *P2ry6*, 5'-CTCACCT-GCATTAGCTTCCAG-3' and 5'-ACAGACTCCACACTACCC-3'; *P2ry12*, 5'-GTTCTCTGGGGTTGATAACCAT-3' and 5'-GCCAGATGACAACAG-AAAGA-3'; *P2ry13*, 5'-CGTTCAGGAAACCTTTGTCA-3' and 5'-

ACACTTCTTCACGGATGATGG-3'. The *Tnf*, *Il22* and *Il23a* primer and probe (FAM MGB Probe) were purchased from Applied Biosystems.

ELISA for faecal IgA. Faeces (50–100 mg) were collected from individual mice, weighed and homogenized in 100 μ l of PBS containing 0.01% NaN_3 per 10 mg of faeces. After centrifugation at 9,000g for 5 min, the supernatants were collected and diluted by 1:1,000. IgA levels were determined by ELISA using a mouse IgA ELISA quantitation kit (Bethyl Laboratories) and TMB solution (eBioscience). Optical densities were determined at a wavelength of 450 nm with a reference wavelength of 570 nm.

T-cell-mediated colitis model. Naive CD4^+ $\text{CD62L}^{\text{high}}$ splenic T cells from Balb/c mice were purified and intraperitoneally transferred into SCID mice (3×10^5 cells per mouse). The mice were then injected with $\alpha\beta$ -ATP in RPMI (1.5 mg per mouse) or medium alone every 3 days for 4 weeks. After 8 weeks, the colons were examined for the numbers of IL-17- and IFN- γ -producing CD4^+ cells or analysed histologically after haematoxylin and eosin staining.

31. Yamamoto, M. *et al.* Role of adaptor TRIF in the MyD88-independent Toll-like receptor signaling pathway. *Science* **301**, 640–643 (2003).
32. Mizoguchi, E. *et al.* Colonic epithelial functional phenotype varies with type and phase of experimental colitis. *Gastroenterology* **125**, 148–161 (2003).

The deubiquitinylation and localization of PTEN are regulated by a HAUSP–PML network

Min Sup Song¹, Leonardo Salmena^{1*}, Arkaitz Carracedo^{1*}, Ainara Egia¹, Francesco Lo-Coco², Julie Teruya-Feldstein³ & Pier Paolo Pandolfi¹

Nuclear exclusion of the PTEN (phosphatase and tensin homologue deleted in chromosome 10) tumour suppressor has been associated with cancer progression^{1–6}. However, the mechanisms leading to this aberrant PTEN localization in human cancers are currently unknown. We have previously reported that ubiquitinylation of PTEN at specific lysine residues regulates its nuclear–cytoplasmic partitioning⁷. Here we show that functional promyelocytic leukaemia protein (PML) nuclear bodies co-ordinate PTEN localization by opposing the action of a previously unknown PTEN-deubiquitinating enzyme, herpesvirus-associated ubiquitin-specific protease (HAUSP, also known as USP7), and that the integrity of this molecular framework is required for PTEN to be able to enter the nucleus. We find that PTEN is aberrantly localized in acute promyelocytic leukaemia, in which PML function is disrupted by the PML–RAR α fusion oncoprotein. Remarkably, treatment with drugs that trigger PML–RAR α degradation, such as all-*trans* retinoic acid or arsenic trioxide, restore nuclear PTEN. We demonstrate that PML opposes the activity of HAUSP towards PTEN through a mechanism involving the adaptor protein DAXX (death domain-associated protein). In support of this paradigm, we show that HAUSP is overexpressed in human prostate cancer and is associated with PTEN nuclear exclusion. Thus, our results delineate a previously unknown PML–DAXX–HAUSP molecular network controlling PTEN deubiquitinylation and trafficking, which is perturbed by oncogenic cues in human cancer, in turn defining a new deubiquitinylation-dependent model for PTEN subcellular compartmentalization.

Although originally thought to function solely in the cytoplasm, there is now compelling evidence that nuclear PTEN is also essential for tumour suppression^{7–11}. To our surprise, we observed that acute promyelocytic leukaemia (APL) biopsies harbouring the t(15;17) chromosomal translocation, encoding the PML–RAR α fusion protein, exhibited a predominantly cytoplasmic PTEN distribution (Fig. 1a and Supplementary Fig. 1; see Methods and Supplementary Fig. 2 for PTEN antibody specificity). APL is commonly treated with all-*trans* retinoic acid (ATRA), an agent that induces degradation of PML–RAR α and restores PML nuclear bodies (Fig. 1b (inset) and Supplementary Fig. 3c, 4a and 5a, b and ref. 12). Examination of APL biopsies and NB4 cells before and after ATRA treatment demonstrated that ATRA effectively restored nuclear PTEN (Fig. 1b, c and Supplementary Figs 3a, b and 4a) with no effect on total PTEN levels (Supplementary Fig. 5b, c). Like ATRA, arsenic trioxide (ATO) is a common therapy for APL, causing a biphasic effect on PML nuclear bodies; short-term treatment promotes the selective proteasome-dependent degradation of PML–RAR α fusion protein and recovery of nuclear bodies, whereas long

exposure induces the degradation of PML and the consequent disassembly of the nuclear bodies^{13–15}. Indeed, this was observed on treatment of NB4 cells with this ATO; specifically, 4 h of treatment reconstituted nuclear bodies and restored PTEN localization to the nucleus, whereas the nuclear bodies were abolished at 24 h of treatment with a concomitant cytoplasmic PTEN localization (Fig. 1c, d and Supplementary Figs 4 and 5), with no significant change of PTEN protein levels (Supplementary Fig. 5b). Furthermore, ATRA did not relocalize PTEN to the nucleus in ATRA-resistant NB4 (NB4R) cells, whereas ATO could control PTEN localization in NB4R as well as in ATRA-sensitive cells (Fig. 1d and Supplementary Fig. 4b).

To validate our findings, we used the MRP8–PML–RAR α (transgenic) mouse model of APL, in which Pten is predominantly cytoplasmic in pre-neoplastic compared to wild-type promyelocytes (Fig. 1e). Notably, ATRA restored nuclear localization of Pten in transgenic promyelocytes (Fig. 1e).

To demonstrate the specificity of PML–RAR α on Pten, *Pml*^{+/-} mouse embryonic fibroblasts (MEFs) were transduced with a retrovirus expressing PML–RAR α . Indeed PML–RAR α was effective at driving nuclear exclusion of Pten (Supplementary Fig. 6a, b), and ATRA treatment restored nuclear Pten (Supplementary Fig. 6c, d). This demonstrates that disassembly of the PML nuclear bodies by PML–RAR α can alter Pten localization. Because Pml is essential for proper nuclear body formation¹⁶, we analysed the localization of Pten in wild-type and *Pml*-null MEFs. In *Pml*-null MEFs, the fraction of cytoplasmic Pten was significantly increased (Fig. 1f and Supplementary Fig. 6e, f) and accompanied by a slight reduction of Pten expression levels (Supplementary Fig. 6f, right panel). Notably, ATRA treatment of NB4 cells induced an enrichment of PTEN in well-defined nuclear structures that co-localized with PML (Supplementary Fig. 7a). Moreover, PTEN co-localized to nuclear bodies in wild-type, but not *Pml*-null, MEFs (Supplementary Fig. 7b). These results are the first, to our knowledge, to highlight a role for PML and PML nuclear bodies in the regulation of PTEN localization and imply that disrupted PTEN localization may have relevance in malignancies in which PML and PML nuclear bodies are compromised as in APL.

Similar to other tumour suppressors such as FOXO4 and p53, PTEN localization can be modulated by ubiquitinylation¹⁷. Strikingly, ATRA treatment increased Pten monoubiquitinylation in NB4 cells (Fig. 2a). Moreover, cells lacking *Pml* displayed reduced levels of Pten monoubiquitinylation (Supplementary Fig. 7c) and reduced nuclear Pten (Fig. 1f and Supplementary Fig. 6e).

To identify the means by which PML nuclear bodies regulate PTEN ubiquitinylation, we focused our attention on nuclear-body-associated proteins that may modulate ubiquitinylation. HAUSP has

¹Cancer Genetics Program, Beth Israel Deaconess Cancer Center and Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA. ²Department of Biopathology, University of Tor Vergata, Rome 00133, Italy. ³Department of Pathology, Memorial Sloan-Kettering Cancer Center, Sloan-Kettering Institute, 1275 York Avenue, New York, New York 10021, USA.

*These authors contributed equally to this work.

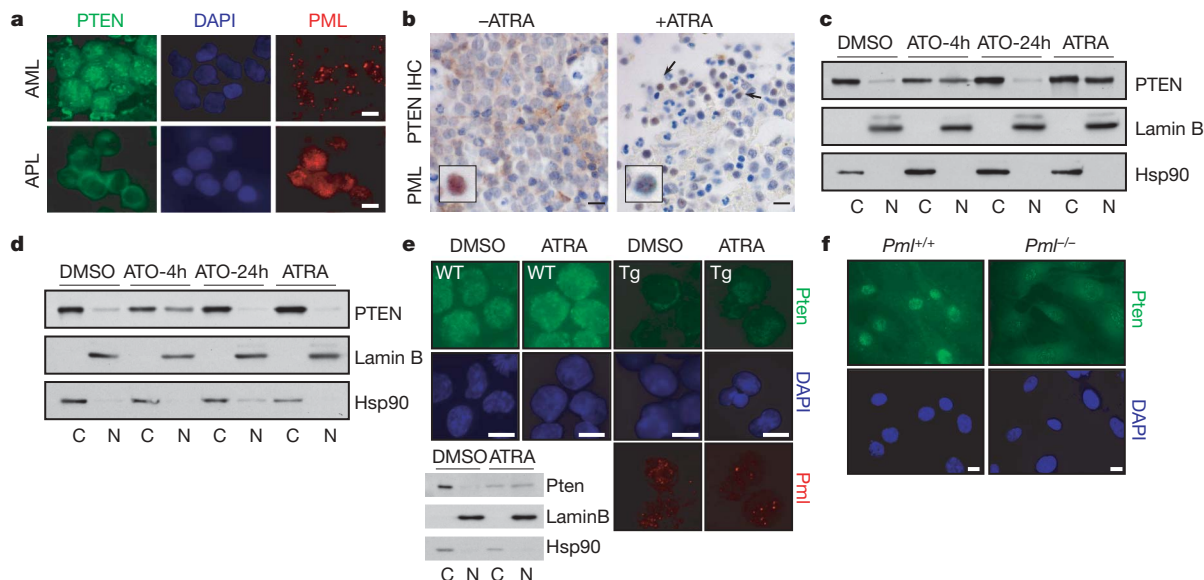


Figure 1 | Aberrant localization of PTEN in APL and *Pml*-null MEFs.

a, Immunofluorescence of PTEN and PML in acute myeloid leukaemia (AML) and APL bone marrow. DAPI, 4,6-diamidino-2-phenylindole. Scale bars, 10 μ m. **b**, Immunohistochemistry (IHC) of PTEN and PML in APL biopsies before and after ATRA treatment. Inset indicates PML localization. Scale bars, 20 μ m. **c**, **d**, Nuclear (N) and cytoplasmic (C) fractionation of

PTEN in NB4 (**c**) or NB4R (**d**) cells after ATO or ATRA treatment. Lamin B and Hsp90 serve as controls. **e**, Immunofluorescence of Pten and Pml in promyelocytes of wild-type (WT) and *MRP8-PML-RAR α* transgenic (Tg) mice after ATRA treatment and fractionation of Pten in transgenic mice (inset). Scale bars, 10 μ m. **f**, Immunofluorescence of Pten in *Pml*^{+/+} and *Pml*^{-/-} MEFs. Scale bars, 10 μ m.

been reported to co-localize to the nuclear bodies, and to critically regulate protein localization through ubiquitinylation¹⁸. Indeed, HAusp localized to Pml nuclear bodies (Supplementary Fig. 7d) and co-immunoprecipitated with Pml (Supplementary Fig. 7e). Moreover, ATRA treatment re-localized HAUSP to PML nuclear bodies (Supplementary Fig. 7f).

Next, by examining the physical and functional interactions between PTEN and HAUSP, we found that they interact directly (Fig. 2b, c and Supplementary Fig. 8a). Importantly, through its enzymatic activity, HAUSP reduced PTEN monoubiquitinylation (Fig. 2d and Supplementary Fig. 8b), because the catalytically inactive (CS) HAUSP mutant, which can bind to PTEN (Supplementary

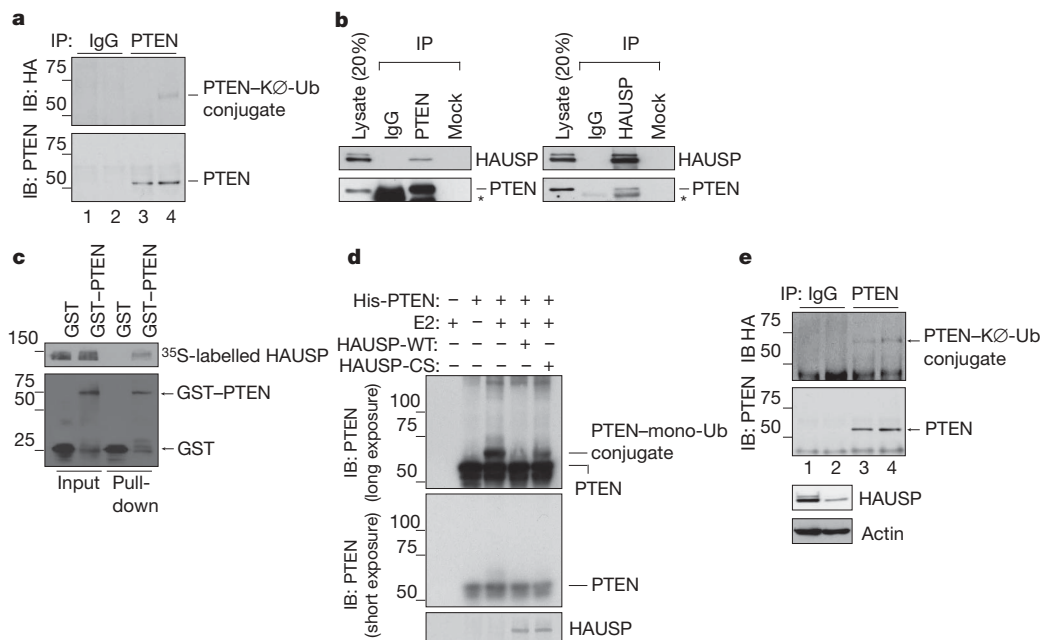


Figure 2 | HAUSP interacts with and deubiquitinates PTEN. **a**, Lysates from haemagglutinin (HA)-K \emptyset -Ub-transfected NB4 cells treated with DMSO (lanes 1 and 3) or ATRA (lanes 2 and 4) were immunoprecipitated (IP) with IgG (lanes 1 and 2) or anti-PTEN (lanes 3 and 4), and analysed for monoubiquitinylation with anti-HA. MW, kDa. **b**, Immunoblotting (IB) of U2OS lysates after IP with the indicated antibodies. Asterisks indicate the heavy chain. **c**, Pull-down of glutathione S-transferase (GST)-PTEN with

³⁵S-labelled *in vitro*-translated HAUSP. **d**, PTEN was incubated with recombinant wild-type (WT) or mutant (CS) HAUSP. **e**, Lysates from HA-K \emptyset -Ub and *Luciferase* siRNA-contransfected (lanes 1 and 3) or *HAUSP* siRNA-cotransfected (lanes 2 and 4) U2OS cells were IP with IgG (lanes 1 and 2) or anti-PTEN (lanes 3 and 4), and analysed for monoubiquitinylation with anti-HA.

Fig. 8c), did not reduce the levels of PTEN monoubiquitinylation. Additionally, RNA interference (RNAi) knockdown of HAUSP increased PTEN monoubiquitinylation (Fig. 2e and Supplementary Fig. 8d). Therefore, these findings uncover HAUSP as the first *bona fide* PTEN deubiquitinylase (DUB).

Monoubiquitinylation has been suggested to be a 'key' for proteins to gain access into other cellular compartments^{19–21}. We found that HAUSP overexpression resulted in the nuclear exclusion of PTEN (Fig. 3a, b and Supplementary Fig. 9a, b). Notably, forced monoubiquitinylation of PTEN with lysine-free ubiquitin (K \emptyset -Ub), which displays mostly nuclear localization, was rendered predominantly cytoplasmic on HAUSP overexpression (Fig. 3c and Supplementary Fig. 9c). Conversely, RNAi knockdown of HAUSP rendered PTEN mainly nuclear (Fig. 3d, e and Supplementary Fig. 9d), in agreement with the notion that monoubiquitinated PTEN is nuclear (Fig. 2d, e, Supplementary Fig. 8b, d and ref. 7). Furthermore, in HAUSP-null HCT116 colon carcinoma cells²², PTEN localization was predominantly nuclear compared to that in parental HCT116 cells (Fig. 3f, g).

Because Lys 289 and Lys 13 are the major residues for PTEN ubiquitinylation⁷, we used cytoplasmic green fluorescent protein (GFP)–PTEN mutants to examine HAUSP activity on these sites (Supplementary Figs 10 and 11). Surprisingly, HAUSP silencing or deletion rendered PTEN(K289E) and PTEN(K13E), but not PTEN(K13,K289E), predominantly nuclear (Supplementary Figs 10a, b and 11a, d). In agreement, co-expression with K \emptyset -Ub promoted PTEN(K289E) and PTEN(K13E), but not PTEN(K13,K289E), nuclear localization (Supplementary Figs 10c, d and 11b, c, e, f). HAUSP overexpression re-localized PTEN(K289E) and PTEN(K13E) to the cytoplasm, suggesting that monoubiquitinylation/deubiquitinylation of one of either site is sufficient to modulate PTEN compartmentalization. (Supplementary Figs 10c, d and 11b, c).

Nuclear exclusion of PTEN has been associated with more aggressive or late-stage cancers, suggesting that cytoplasmic PTEN lacks full tumour suppressive function^{1–6}. To address the role of nuclear PTEN in tumour suppression, we expressed GFP-tagged wild-type PTEN or PTEN(K13,K289E) in (PTEN-null) PC3 prostate cancer cells and analysed apoptosis by TUNEL and annexin V (Supplementary Fig. 11g and data not shown). As expected, PTEN expression increased apoptosis relative to control cells, which was further increased by its forced monoubiquitinylation. However, PTEN(K13,K289E) overexpression resulted in a lower apoptotic index and did not increase the apoptosis on co-expression of K \emptyset -Ub. These results support the notion that nuclear-excluded PTEN has a decreased apoptotic potential, and therefore less tumour suppressive ability. These findings are of great relevance because the HAUSP/PTEN regulatory network is frequently aberrant in human prostate cancer (see below).

Our results demonstrate that deubiquitinylation of PTEN by HAUSP has a crucial role in PTEN localization and function, which implies that alterations in the levels of PTEN DUBs may account for the nuclear exclusion of PTEN observed in many cancers. Indeed, we found that HAUSP is overexpressed in prostate cancer, and, more importantly, high levels of HAUSP are directly correlated with tumour aggressiveness (Supplementary Fig. 12). Moreover, the association between HAUSP levels and PTEN localization indicated the existence of a direct relationship ($P < 0.000008$; χ^2 -test) between HAUSP levels and nuclear exclusion of PTEN (Fig. 3h). These results are the first to report that HAUSP is overexpressed in cancer and directly associated with tumour aggressiveness. Furthermore, the correlation with PTEN localization lends support to the notion that overexpression of this DUB impacts on the correct localization of the tumour suppressor PTEN, and therefore on tumorigenesis.

Collectively, our findings led us to hypothesize that PML and HAUSP exert opposite functions with respect to PTEN monoubiquitinylation and localization. Indeed, the ability of HAUSP overexpression to localize PTEN to the cytoplasm was markedly inhibited by co-overexpression of PML, such that PTEN was predominantly

localized to the nucleus (Fig. 4a) and partially co-localized with PML and HAUSP (Supplementary Fig. 13a). Furthermore, PML, but not PML–RAR α , was able to inhibit the deubiquitinylation of PTEN observed on HAUSP overexpression alone (Fig. 4b and Supplementary Fig. 13b). Similarly, *Hausp* silencing in *Pml*-null MEFs restored a wild-type-like distribution of Pten (Fig. 4c and Supplementary Fig. 13c). Overall, we have identified a previously unknown pathway by which PML regulates PTEN compartmentalization through the inhibition of HAUSP-mediated deubiquitinylation.

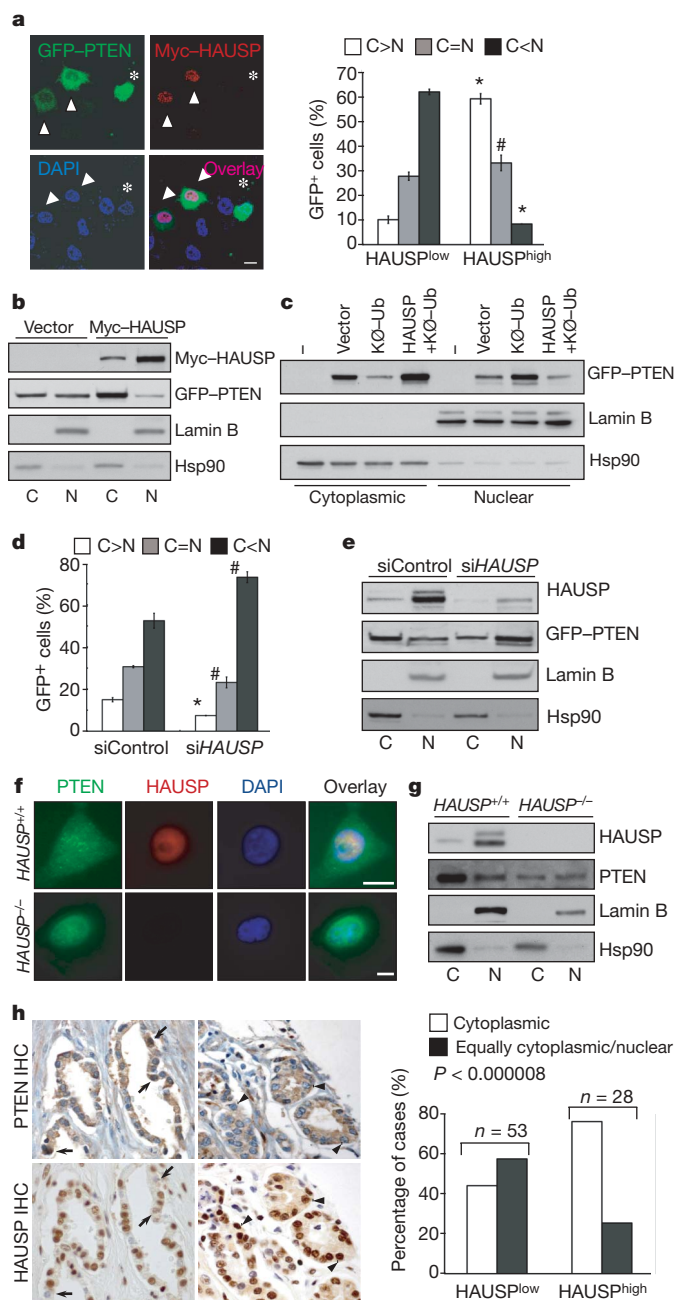


Figure 3 | HAUSP regulates PTEN localization. **a, b**, Immunofluorescence (**a**) and fractionation (**b**) in control or Myc–HAUSP-transfected PC3 cells. Arrowhead, HAUSP^{high}; asterisk, HAUSP^{low}. Scale bar, 10 μ m. Error bars, s.d. **c**, Fractionation in HA–K \emptyset -Ub or/and Myc–HAUSP-transfected PC3 cells. **d, e**, Immunofluorescence (**d**) and fractionation (**e**) in PC3 cells on control or HAUSP siRNA transfection. Error bars, s.d. **f, g**, Immunofluorescence (**f**) and fractionation (**g**) of PTEN in HAUSP^{+/+} and HAUSP^{-/-} HCT116 cells. Scale bars, 10 μ m. **h**, IHC of prostate TMA for PTEN and HAUSP (left) and quantification (right, $n = 81$). Arrow, nuclear PTEN (HAUSP^{low}); arrowhead, cytoplasmic PTEN (HAUSP^{high}).

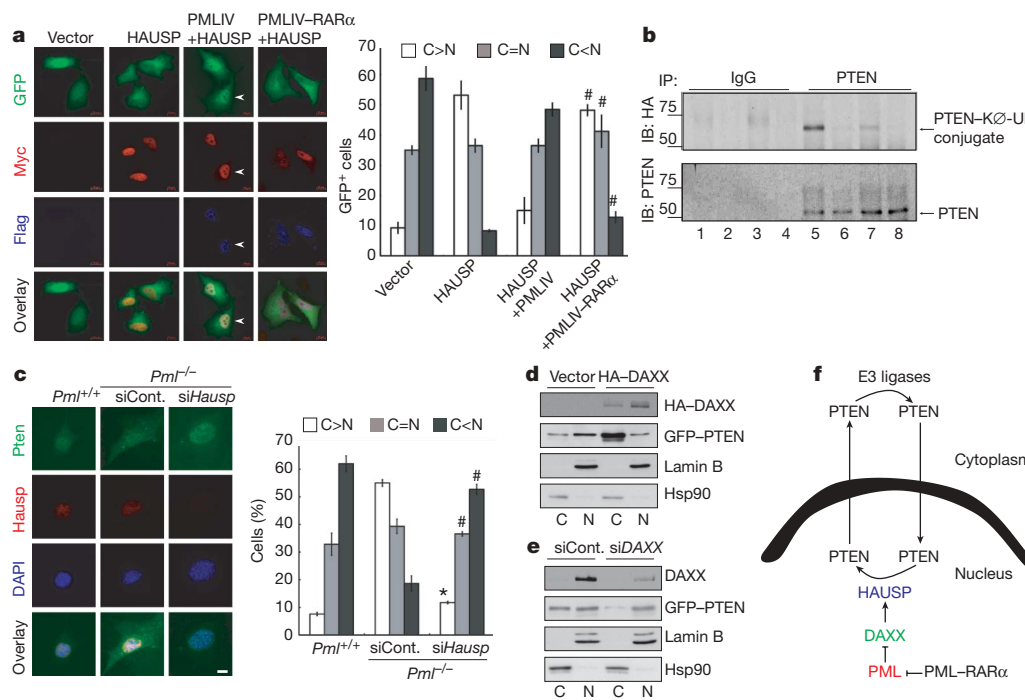


Figure 4 | PML opposes HAUSP-mediated PTEN deubiquitylation.

a, Immunofluorescence of GFP-PTEN in Myc-HAUSP-, Flag-PMLIV- and PMLIV-RAR α -transfected PC3 cells. Error bars, s.d. **b**, HA-Ub IB in SV40-*Pml*^{-/-} MEFs cotransfected with HA-K \emptyset -Ub and Flag-PTEN (lanes 1 and 5), Flag-PTEN and Myc-HAUSP (lanes 2 and 6), Flag-PTEN, Myc-HAUSP and Flag-PMLIV (lanes 3 and 7), or Flag-PTEN, Myc-HAUSP and Flag-PMLIV-RAR α (lanes 4 and 8). Lanes 1–4 for PTEN and 5–8 for IgG IP.

To elucidate the mechanism by which PML opposes HAUSP function towards PTEN, we examined the consequences of expression of three different PML-mutant proteins on PTEN localization (Supplementary Fig. 14a). Surprisingly, sumoylation-deficient PML species, which have been shown to form non-functional nuclear bodies²³, were unable to localize PTEN to the nucleus compared to the wild-type control and a PML RING domain mutant (PML-CS; C57S,C60S, Supplementary Fig. 14b). This suggests that the presence of functional nuclear bodies is requisite for HAUSP function. On this basis, we focused on DAXX, a nuclear-body-residing protein that has been demonstrated to modulate HAUSP function^{24,25} and is not repressed by the sumoylation-defective PML mutants²⁶. Indeed, DAXX overexpression led to cytoplasmic PTEN localization (Fig. 4d and Supplementary Fig. 15a) whereas DAXX silencing increased the nuclear fraction of PTEN (Fig. 4e and Supplementary Fig. 15b).

Taken together, our findings have allowed us to reach a number of relevant conclusions. First, we have identified HAUSP as a critical and essential enzyme for PTEN deubiquitylation and subcellular localization. Nuclear exclusion of PTEN is associated with a more aggressive cancers^{1,7}. Also, PTEN has been proposed to regulate genomic stability from the nucleus⁹, and it can oppose anchorage-independent growth²⁷. In line with this idea, our data show that monoubiquitylated PTEN, which is mainly nuclear, possesses greater apoptotic potential. This suggests that the tumour suppressive activity of PTEN is at least in part due to its ability to reside in the nucleus. Until now, only the E3-ligase NEDD4-1 has been reported to ubiquitylate PTEN²⁸, although additional E3 ligases certainly serve this purpose. Consequently, the notion of ubiquitin-dependency for localization implies that other enzymes capable of deubiquitylating PTEN may also have an important role in PTEN localization. Therefore, the identification of HAUSP as a critical DUB for PTEN subcellular localization lends strong support to the notion that cellular localization of PTEN is controlled by its ubiquitylation status.

c, Immunofluorescence of Pten in control or *HAusp* siRNA-transfected primary *Pml*^{-/-} MEFs (**P* < 0.01; #*P* < 0.05). Scale bars, 10 μ m. An average (%) of 300 cells was counted for quantification of Pten localization in three independent experiments. Error bars, s.d. **d, e**, Fractionation in control or HA-DAXX (**d**) and siRNA control or DAXX-transfected (**e**) PC3 cells. **f**, A model for the control of PTEN monoubiquitylation and localization by HAUSP, DAXX and PML.

Second, we propose a new model for PTEN trafficking (Fig. 4f) whereby in physiological conditions PTEN ubiquitylation by E3-ligases allows PTEN to accumulate in the nucleus, whereas PTEN deubiquitylation by HAUSP favours PTEN accumulation in the cytoplasm. According to this model, PML has a critical regulatory role by inhibiting HAUSP activity through DAXX, which in turn favours PTEN nuclear localization.

Finally, these findings have important implications for cancer because the new regulatory network we have identified can be perturbed at multiple levels in oncogenic conditions (Fig. 4f). First, HAUSP overexpression leads to a predominantly nuclear-excluded PTEN. The fact that this phenotype is associated with more aggressive prostate cancer implies that HAUSP may behave as an oncogene when aberrantly expressed through its ability to disrupt PTEN function through delocalization. Second, this network includes the tumour suppressor PML, suggesting that PTEN delocalization may be an important consequence in *PML*-loss-associated cancers¹⁹. Third, as PTEN is excluded from the nucleus of APL blasts, loss of PTEN nuclear function could have a role in the pathogenesis of this and possibly other forms of leukaemia. The identification of such a network and its deregulated hubs in human cancers will ultimately allow the development of drugs that restore normal PTEN localization towards effective therapy, as epitomized by ATRA and ATO treatment in APL cells.

METHODS SUMMARY

Cell culture. All MEFs were prepared from embryos at day 13.5 of development (E13.5). Early passage (P2–P5) MEFs or SV40 immortalized MEFs were used in all experiments, except as indicated. NB4 cells, a human APL cell line bearing the t(15;17), and ATRA-resistant NB4 (NB4R) cells were a gift from M. Lanotte. *HAUSP*^{+/+} and *HAUSP*^{-/-} HCT116 cells were provided by B. Vogelstein. ATRA (Sigma) and ATO (Sigma) were prepared at a concentration of 1 mmol l⁻¹ in PBS and 10 mol l⁻¹ in NaOH, respectively. Promyelocytes were isolated by flushing the femur and tibia bone marrow and collected by magnetic separation with anti-CD11b and anti-Gr-1.

Antibodies. The following antibodies were used: PTEN (for IF/IHC, 6H2.1, Cascade Bioscience; for immunoblot, Cell Signaling), PML (Ab1370, MAB3738, Chemicon; H-238, Santa Cruz;), HAUSP (Bethyl Laboratories), Myc (Cell Signaling), Flag (Sigma), DAXX (Santa Cruz), lamin B1 (Abcam), Hsp90 (BD Transduction Laboratories), RAR α (Affinity Bioreagents), β -actin (Sigma), glutathione S-transferase (Santa Cruz), haemagglutinin (Covance) and ubiquitin (BD Pharmingen; Biomol International). Immunoblotting, immunofluorescence, co-immunoprecipitation and the *in vitro* binding assay were performed with these antibodies as described⁷.

Immunohistochemistry on patient samples and tumour tissue microarrays. Prostate tumour tissue microarrays were constructed using a fully automated Beecher Instrument, ATA-27. The study cohort comprised APL and prostate tumours consecutively ascertained at the Memorial Sloan-Kettering Cancer Center (MSKCC). All biopsies were evaluated at MSKCC, and the histological diagnosis was based on established criteria.

Nuclear/cytoplasmic fractionation. Nuclear/cytoplasmic fractionation was performed as described⁷. Nuclear to cytoplasmic ratios were calculated using the following densities (PTEN-nuc/lamin B)/(PTEN-cyto/Hsp90).

Ubiquitinylation assays. *In vivo* ubiquitinylation was performed as described⁷. Immunoprecipitates with anti-PTEN were washed twice with 0.5 M LiCl in TBS and twice with TBS, and then resolved by SDS-PAGE. *In vitro* ubiquitinylation was performed as described⁷.

Statistical analysis. Statistical significance was evaluated by using the Student's *t*-test and χ^2 analysis when required.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 December 2007; accepted 31 July 2008.

Published online 20 August 2008.

- Tachibana, M. *et al.* Expression and prognostic significance of PTEN protein in patients with esophageal squamous cell carcinoma. *Cancer* **94**, 1955–1960 (2002).
- Whiteman, D. C. *et al.* Nuclear PTEN expression and clinicopathologic features in a population-based series of primary cutaneous melanoma. *Int. J. Cancer* **99**, 63–67 (2002).
- Zhou, X. P. *et al.* Epigenetic PTEN silencing in malignant melanomas without PTEN mutation. *Am. J. Pathol.* **157**, 1123–1128 (2000).
- Zhou, X. P. *et al.* PTEN mutational spectra, expression levels, and subcellular localization in microsatellite stable and unstable colorectal cancers. *Am. J. Pathol.* **161**, 439–447 (2002).
- Perren, A. *et al.* Mutation and expression analyses reveal differential subcellular compartmentalization of PTEN in endocrine pancreatic tumors compared to normal islet cells. *Am. J. Pathol.* **157**, 1097–1103 (2000).
- Fridberg, M. *et al.* Protein expression and cellular localization in two prognostic subgroups of diffuse large B-cell lymphoma: higher expression of ZAP70 and PKC- β II in the non-germinal center group and poor survival in patients deficient in nuclear PTEN. *Leuk. Lymphoma* **48**, 2221–2232 (2007).
- Trotman, L. C. *et al.* Ubiquitination regulates PTEN nuclear import and tumor suppression. *Cell* **128**, 141–156 (2007).
- Baker, S. J. PTEN enters the nuclear age. *Cell* **128**, 25–28 (2007).
- Shen, W. H. *et al.* Essential role for nuclear PTEN in maintaining chromosomal integrity. *Cell* **128**, 157–170 (2007).
- Carracedo, A., Salmena, L. & Pandolfi, P. P. Snapshot: PTEN signaling pathways. *Cell* **133**, 550 (2008).
- Salmena, L., Carracedo, A. & Pandolfi, P. P. Tenets of PTEN tumor suppression. *Cell* **133**, 403–414 (2008).
- Tallman, M. S. Acute promyelocytic leukemia as a paradigm for targeted therapy. *Semin. Hematol.* **41**, 27–32 (2004).
- Lallemend-Breitenbach, V. *et al.* Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As₂O₃-induced PML or PML/retinoic acid receptor α degradation. *J. Exp. Med.* **193**, 1361–1371 (2001).
- Lallemend-Breitenbach, V. *et al.* Arsenic degrades PML or PML-RAR α through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nature Cell Biol.* **10**, 547–555 (2008).
- Tatham, M. H. *et al.* RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature Cell Biol.* **10**, 538–546 (2008).
- Bernardi, R. & Pandolfi, P. P. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nature Rev. Mol. Cell Biol.* **8**, 1006–1016 (2007).
- Salmena, L. & Pandolfi, P. P. Changing venues for tumour suppression: balancing destruction and localization by monoubiquitylation. *Nature Rev. Cancer* **7**, 409–413 (2007).
- Everett, R. D. *et al.* A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.* **16**, 1519–1530 (1997).
- Li, M. *et al.* Mono- versus polyubiquitination: differential control of p53 fate by Mdm2. *Science* **302**, 1972–1975 (2003).
- van der Horst, A. *et al.* FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP. *Nature Cell Biol.* **8**, 1064–1073 (2006).
- Marchenko, N. D., Wolff, S., Erster, S., Becker, K. & Moll, U. M. Monoubiquitylation promotes mitochondrial p53 translocation. *EMBO J.* **26**, 923–934 (2007).
- Cummins, J. M. *et al.* Tumour suppression: disruption of HAUSP gene stabilizes p53. *Nature* **428**, doi:10.1038/nature02501 (2004).
- Shen, T. H., Lin, H. K., Scaglioni, P. P., Yung, T. M. & Pandolfi, P. P. The mechanisms of PML-nuclear body formation. *Mol. Cell* **24**, 331–339 (2006).
- Song, M. S., Song, S. J., Kim, S. Y., Oh, H. J. & Lim, D. S. The tumour suppressor RASSF1A promotes MDM2 self-ubiquitination by disrupting the MDM2–DAXX–HAUSP complex. *EMBO J.* **27**, 1863–1874 (2008).
- Tang, J. *et al.* Critical role for Daxx in regulating Mdm2. *Nature Cell Biol.* **8**, 855–862 (2006).
- Lin, D. Y. *et al.* Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Mol. Cell* **24**, 341–354 (2006).
- Liu, J. L. *et al.* Nuclear PTEN-mediated growth suppression is independent of Akt down-regulation. *Mol. Cell Biol.* **25**, 6211–6224 (2005).
- Wang, X. *et al.* NEDD4-1 is a proto-oncogenic ubiquitin ligase for PTEN. *Cell* **128**, 129–139 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank B. Vogelstein, K. H. Baek, M. Lanotte and X. Jiang for sharing reagents and W. Gu for critical discussions. We thank all members of the Pandolfi laboratory, in particular K. Ito, S. Majid and L. Polisenio, for technical support, advice and discussion. This work was supported by NIH grants to P.P.P. L.S. is supported by the International Human Frontier Science Program Organization, and A.C. is supported by the European Molecular Biology Organization.

Author Contributions The experiments were conceived and designed by M.S.S., L.S., A.C. and P.P.P. Experiments were performed by M.S.S., L.S., A.C. and A.E. F.L.-C. provided APL and AML samples. J.T.-F. provided and scored IHC of prostate cancer tissue microarray and APL samples. Data were analysed by M.S.S., L.S., A.C., J.T.-F. and P.P.P. The paper was written by M.S.S., L.S., A.C. and P.P.P.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.P.P. (ppandolf@bidmc.harvard.edu).

METHODS

Cell culture. Primary *Pml*^{+/+}, *Pml*^{+/-} and *Pml*^{-/-} MEFs were prepared from embryos at E13.5. Early passage (P2–P5) MEFs or SV40 immortalized MEFs were used in all experiments, except as indicated.

***In vitro* promyelocytes culture.** Bone marrow cells from wild-type (*n* = 5) and *MRP8-PML-RARα* transgenic (*n* = 5) mice were obtained by flushing the femur and tibia bone marrow in RPMI containing 15% FBS. Promyelocytes were collected using the MACS separation column (Miltenyi Biotec) after incubation with anti-CD11b (M1170) and anti-Gr-1 (Rb6-8c5, eBioscience), cultured in RPMI and then treated with DMSO or ATRA for 72 h.

Antibodies. The following antibodies were used for immunofluorescence analysis: PTEN (1:200, 6H2.1, Cascade Bioscience; 1:200, 138G6, Cell Signaling; 1:200, 10P03, Neomarkers; 1:500, polyclonal, R&D Systems), PML (1:500, Ab1370, Chemicon; 1:200, H-238, Santa Cruz; 1:1,000, MAB3738, Chemicon), HAUSP (1:500, Bethyl Laboratories), Myc (1:1,000, Cell Signaling), Flag (1:1,000, Sigma) and DAXX (1:200, Santa Cruz). We used the following antibodies for immunohistochemistry analysis: PTEN (1:100, Cascade), PML (1:200, Santa Cruz; 1:200, Chemicon) and HAUSP (1:200, Bethyl Laboratories). The following antibodies were used for immunoblot analysis: PTEN (1:1,000, Cell Signaling; 1:1,000, Cascade Bioscience), lamin B1 (1:2,000, Abcam), Hsp90 (1:2,000, BD Transduction Laboratories), PML (1:1,000, MAB3738, Chemicon; 1:1,000, Santa Cruz), *RARα* (1:1,000, Affinity Bioreagents), HAUSP (1:2,000, Bethyl Laboratories), β -actin (1:5,000, Sigma), Myc (1:2,000, Cell Signaling), Flag (1:2,000, Sigma), GST (1:5,000, Santa Cruz), HA (1:2,000, Covance) and ubiquitin (1:1,000, BD Pharmingen; 1:1,000, Biomol International).

Quantitative real-time PCR. RNA was isolated with the RNeasy Protect kit (Qiagen) and including a DNase digestion step using the RNase-free DNase kit (Qiagen). Complementary DNA was obtained with Transcriptor (Roche). Taqman probes were obtained from Applied Biosystems. Amplifications were run in a 7900 Real-Time PCR System (Applied Biosystems). Each value was adjusted by using β -glucuronidase levels as a reference.

***In vitro* promyelocytes culture.** Bone marrow cells from wild-type mice (*n* = 5) and *MRP8-PML-RARα* transgenic mice (*n* = 5) were obtained by flushing the femur and tibia bone marrow in RPMI containing 15% FBS. Promyelocytes were collected using the MACS separation column (Miltenyi Biotec) after incubation with anti-CD11b (M1170) and anti-Gr-1 (Rb6-8c5, eBioscience), cultured in RPMI and then treated with DMSO or ATRA for 72 h.

Immunohistochemistry on patient samples and tumour tissue microarrays. Bone marrow smear slides from nine patients with AML and nine patients with APL were analysed by immunofluorescence with anti-PTEN (6H2.1, Cascade) and PML (Ab1370, Chemicon) as previously described⁷. APL patient sample

slides before (*n* = 40) and after (*n* = 9) ATRA treatment were also stained for PTEN and PML.

Plasmids and recombinant proteins. pEGFP-PTEN, pEGFP-PTEN(K289E), -PTEN(K13E), -PTEN(K13;289E), pCMV-Tag2B-PTEN, pBabe-PMLIV, pBabe-PMLIV-RAR α , pCMV-Tag2B-PMLIV wild-type, -PMLIV-CS, -PMLIV-3mas -PML-3mds, -PMLIV-RAR α , and pCMV-HA-Ub were described previously^{7,23,29,30}. HA-K \emptyset -Ub and HA-DAXX were also described previously^{7,23}. Recombinant GST-PTEN and His-PTEN protein were provided by X. Jiang²⁸. Myc-HAUSP wild type and the mutant form (CS) were a gift from K. H. Baek. For purification of recombinant HAUSP, Flag-HAUSP-WT and -CS were expressed in 293T cells and purified on M2 column.

siRNA transfection. SMARTpool HAUSP (Dharmacon) represents four pooled SMART-selected siRNA duplexes that target HAUSP (sense sequences are: CUAAGGACCCUGCAAAUUAUU; GUGGUUACGUUAUCAAUUAUU; UGA CGUGUCUCUUGAUAAAUU; GAAGGUACUUUAAGAGAUUU). Cells were transfected with 200 nM of either luciferase or scrambled siRNA (controls) or HAUSP siRNA using the DharmaFECT transfection reagent (Thermo Scientific). Twenty-four hours after transfection, cells were plated for additional transfection for immunofluorescence or immunoblotting.

Immunoblotting, immunoprecipitation and *in vitro* binding assay. Immunoblotting, (co-)immunoprecipitation and the *in vitro* binding assay were performed as described⁷. The following antibodies were used for immunoblotting: PTEN (1:1,000, Cell Signaling; 1:1,000, Cascade Bioscience), lamin B (1:2,000, Abcam), Hsp90 (1:2,000, BD Transduction Laboratories), actin (1:5,000, Sigma), Myc (1:2,000, Cell Signaling), Flag (1:2,000, Sigma), GST (1:5,000, Santa Cruz), HA (1:2,000, Covance), HAUSP (1:2,000, Bethyl Laboratories) and DAXX (1:5,000, Epitomics).

Immunofluorescence. Immunofluorescence was performed as described⁷. Secondary antibodies used were: Alexa Fluor-488 or -594 (Molecular Probes) or AMCA-conjugated IgG (Jackson ImmunoResearch Laboratories). Coverslips were analysed by an inverted microscope (Zeiss Axiovert 200) or a confocal laser-scanning microscope (Leica TSC STED). Data were processed with Adobe Photoshop 7.0 software.

Apoptosis assay. PC3 cells were grown on coverslips and transfected with the indicated plasmids. Forty-eight hours after the transfection, cells were fixed in 4% paraformaldehyde and stained for TUNEL (Roche) following the manufacturer's instructions. Positive cells were quantified in Zeiss Axiovert 200 inverted microscope.

29. Bernardi, R. *et al.* PML regulates p53 stability by sequestering Mdm2 to the nucleolus. *Nature Cell Biol.* **6**, 665–672 (2004).

30. Scaglioni, P. P. *et al.* A CK2-dependent mechanism for degradation of the PML tumor suppressor. *Cell* **126**, 269–283 (2006).

LETTERS

Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism

Kyohei Arita¹, Mariko Ariyoshi¹, Hidehito Tochio^{1,2}, Yusuke Nakamura³ & Masahiro Shirakawa^{1,2,4}

DNA methylation of CpG dinucleotides is an important epigenetic modification of mammalian genomes and is essential for the regulation of chromatin structure, of gene expression and of genome stability^{1,2}. Differences in DNA methylation patterns underlie a wide range of biological processes, such as genomic imprinting, inactivation of the X chromosome, embryogenesis, and carcinogenesis^{3–6}. Inheritance of the epigenetic methylation pattern is mediated by the enzyme DNA methyltransferase 1 (Dnmt1), which methylates newly synthesized CpG sequences during DNA replication, depending on the methylation status of the template strands^{7,8}. The protein UHRF1 (also known as Np95 and ICBP90) recognizes hemi-methylation sites via a SET and RING-associated (SRA) domain and directs Dnmt1 to these sites^{9–11}. Here we report the crystal structures of the SRA domain in free and hemi-methylated DNA-bound states. The SRA domain folds into a globular structure with a basic concave surface formed by highly conserved residues. Binding of DNA to the concave surface causes a loop and an amino-terminal tail of the SRA domain to fold into DNA interfaces at the major and minor grooves of the methylation site. In contrast to fully methylated CpG sites recognized by the methyl-CpG-binding domain^{12,13}, the methylcytosine base at the hemi-methylated site is flipped out of the DNA helix in the SRA–DNA complex and fits tightly into a protein pocket on the concave surface. The complex structure suggests that the successive flip out of the pre-existing methylated cytosine and the target cytosine to be methylated is associated with the coordinated transfer of the hemi-methylated CpG site from UHRF1 to Dnmt1.

The crystal structure of the unliganded SRA domain (amino-acid residues 405–613) of mouse UHRF1 was determined at 1.77 Å resolution (see Methods). Subsequently, co-crystallization was performed with two different 12-base-pair (bp) oligonucleotide duplexes (Supplementary Fig. 1a), each containing a hemi-methylated CpG site. We determined three crystal structures of the complexes, at 1.6, 2.55 and 2.6 Å resolution (see Methods). The structures are essentially identical in both the protein and DNA portions (see Supplementary Results). Only the 1.6-Å resolution structure is described below.

The SRA domain is folded into a single globular structure consisting of a five-stranded mixed β -sheet, comprising strands β 1, β 2, β 3, β 5 and β 6, and a three-stranded anti-parallel β -sheet, comprising strands β 4, β 7 and β 6, and their associated four helices, α 1– α 4 (Fig. 1a and Supplementary Fig. 2). The two β -sheets are connected by a long arch-shaped β 6 strand so that they partly stack one on top of the other and form a twisted shell-like β -sandwich. The overall shape of the domain looks like a saddle, having a concave surface at the open side of the shell-like β -sandwich. The SRA-domain fold generates a well-defined hydrophobic core formed by residues with high

degrees of conservation (Supplementary Fig. 3). Mapping of conserved residues on the molecular surface of the SRA domain indicates that they contribute to a positively charged patch that covers a large portion of the inner side of the concave surface (Fig. 1b and Supplementary Fig. 4a). Within the patch, a deep pocket was present, which is formed by residues well conserved among SRA domains.

Hemi-methylated DNA fits into the conserved concave surface of the SRA domain in the complex structure (Fig. 2a and Supplementary Fig. 4b). Binding of DNA induces structural changes in the N-terminal 7 residues (405–411), and in 11 residues in a loop region (489–499), designated the N-tail and finger loop, respectively, which hold DNA at the hemi-methylation site from both sides (Supplementary Fig. 5a). Electron densities for these residues were not observed in the unliganded SRA. The methylated cytosine base is flipped out of the DNA duplex and is inserted into the conserved pocket on the inner concave surface of the SRA domain (Fig. 2a). In contrast, contacts with the bases of the DNA double helix are made exclusively by the two induced loops: the finger loop is inserted into the major groove of the hemi-methylation site, and the N-tail is embedded in the minor groove, adopting a partly helical conformation (Supplementary Fig. 5b, c). Despite the flipped-out base, the DNA in the complex is unbent and essentially retains the regular B-form conformation. Base flipping has been already observed in DNA structures bound to bacterial DNA methyltransferases and in

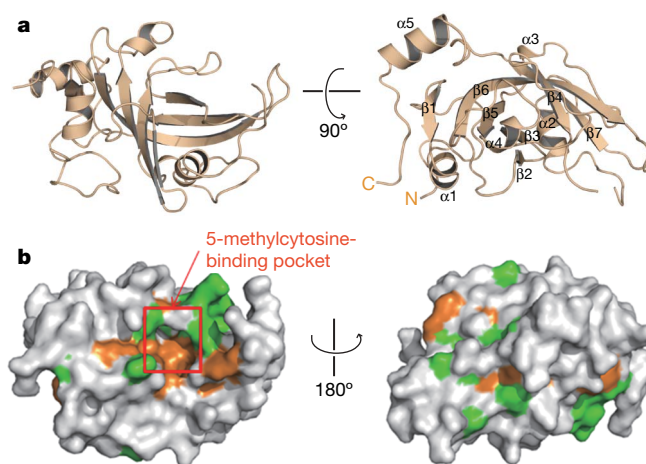


Figure 1 | Overall structure of the SRA domain. **a**, Ribbon representation of the unliganded SRA domain of UHRF1. **b**, Distribution of conserved residues on the molecular surface of the unliganded SRA domain. Amino-acid residues are coloured according to their sequence conservation level: orange, fully conserved; light green, conservative substitutions. In the left panel the molecule is shown in the same orientation as in the left panel of **a**.

¹Graduate School of Engineering, Kyoto University, Kyoto 615-8510, Japan. ²Japan Science and Technology Agency, CREST, 4-1-18, Honcho, Kawaguchi-shi, Saitama 332-0012, Japan. ³Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. ⁴RIKEN, Yokohama Institute, Suehirocho, Tsurumi, Yokohama 230-0045, Japan.

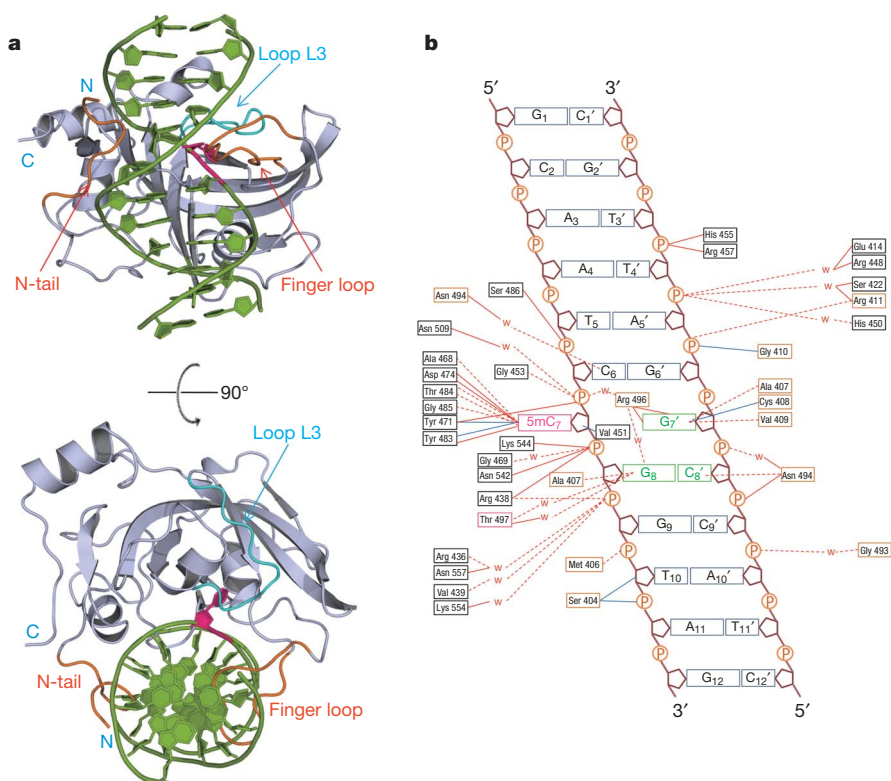


Figure 2 | Overall structure of the SRA-hemi-methylated CpG DNA complex. **a**, Ribbon representation of the SRA–DNA complex in the same orientations as in Fig. 1a: grey, SRA; orange, N-tail and finger loop; light blue, loop L3; light green, DNA; magenta, flipped-out base 5mC₇. **b**, Schematic representation of the protein–DNA interactions observed in

some classes of DNA repair enzymes^{14–18}. All these enzymes have catalytic activities for the modification or elimination of the flipped-out bases or of orphaned bases on the complementary strands. The SRA domain shows no apparent resemblance to these enzymes in either its primary or its tertiary structure.

The flipped-out methylated cytosine (5mC₇) ring is specifically recognized in the conserved pocket (Fig. 3a). The position of the ring is fixed by sandwich-stacking interactions with the aromatic rings of Tyr 471 and Tyr 483. The 5'-methyl group of 5mC₇ is packed against a cavity within the pocket, and is within van der Waals contact distances (less than 4.2 Å) from Thr 471, Tyr 483, Gly 485 and Ser 486 (Supplementary Fig. 6). The methyl group is also positioned within CH–O hydrogen-bond distances (less than 4 Å) from Tyr 483, Thr 484 and Gly 485 of the SRA domain and the oxygen of the backbone phosphate P₇. The discrimination of 5mC₇ from other base rings such as thymine is further accomplished by three hydrogen bonds involving the N4 and O2 atoms of 5mC₇ and SRA residues, and possibly by another bond between the N3 atom and the Asp 474 carboxyl, for which the protonation of either N3 or the carboxyl is required (Fig. 3a). Three of six 5mC₇-contacting residues are contributed by loop L3, which moves about 2.9 Å towards the 5mC₇ base on the binding of DNA so as to establish the tight packing of the base against the pocket (Supplementary Fig. 5d).

Another important consequence of base flipping is the invasion by the SRA finger loop of the vacancy in the double helix, which recognizes both the orphaned guanine, G₇', and the adjacent G₈•C₈' pair containing the target cytosine for the maintenance methylation by Dnmt1. At the tip of the loop, Arg 496 guanidino donates hydrogen bonds to the O6 and N7 atoms of G₇', which is thus held near its intact position in the helix (Fig. 3b). This base also has minor groove contacts through hydrogen bonds between its N2 and the carbonyls of Ala 407 and Val 409 of the N-tail. The adjacent G₈•C₈' base pair is

the crystal structure. Hydrogen bonds between SRA and DNA are shown by red lines: dotted, main-chain contacts; solid, side-chain contacts. Van der Waals contacts between SRA and DNA are shown by blue lines. W, water molecules mediating indirect interactions.

also recognized by both the major and minor groove contacts (Fig. 3c). The G₈ purine ring is stabilized by the stacking interaction with Arg 496, which invades the space after base flipping. The ring then forms water-mediated hydrogen bonds with the amide of Arg 496 and the amide and the O_γ atom of Thr 497 at the major groove and a direct hydrogen bond with the Ala 407 carbonyl at the minor groove. The C₈' of the complementary strand is positioned close to Asn 494 of the finger loop, enabling the N4 atom to form a hydrogen bond with the carbonyl of this residue. This residue is of special importance because it buttresses Arg 496, a key residue for G₇' and G₈ recognition, through hydrogen bonding between their side chains (Fig. 3d). The main-chain carbonyl oxygen and the side-chain C_γ and N_δ of Asn 494 are situated so as to make steric clashes with the methyl group if the fifth position of C₈' is methylated (Fig. 3d). The lower affinity of the SRA domain for DNA containing symmetrical methyl-CpG sites⁹ (Supplementary Fig. 1b) might therefore result from interference between the methyl group of the DNA and Asn 494. Collectively, the extensive specific contacts to the 5mCpG/CpG bases outside and within the helix observed in the complex structure explain the selectivity of the SRA domain for hemi-methylated DNA over unmethylated and fully methylated DNA. Direct base–protein contacts are limited to the 5mCpG/CpG sequence (Fig. 2b), suggesting that SRA binding is essentially independent of the DNA bases surrounding 5mCpG/CpG.

The SRA domain's mode of recognition of hemi-methylated DNA is in striking contrast to that of fully methylated DNA by methyl-DNA-binding domains (MBDs). Fully methylated DNA in complex with MBD has a nearly normal B-form conformation, and both methylcytosine bases remain in the double helix^{12,13}. The contact area of each of the methylcytosine rings with MBD1 is 85 or 112 Å². The summed contact area of MBD to the two methylcytosine bases at the fully methylated site is comparable to that of the SRA domain to the

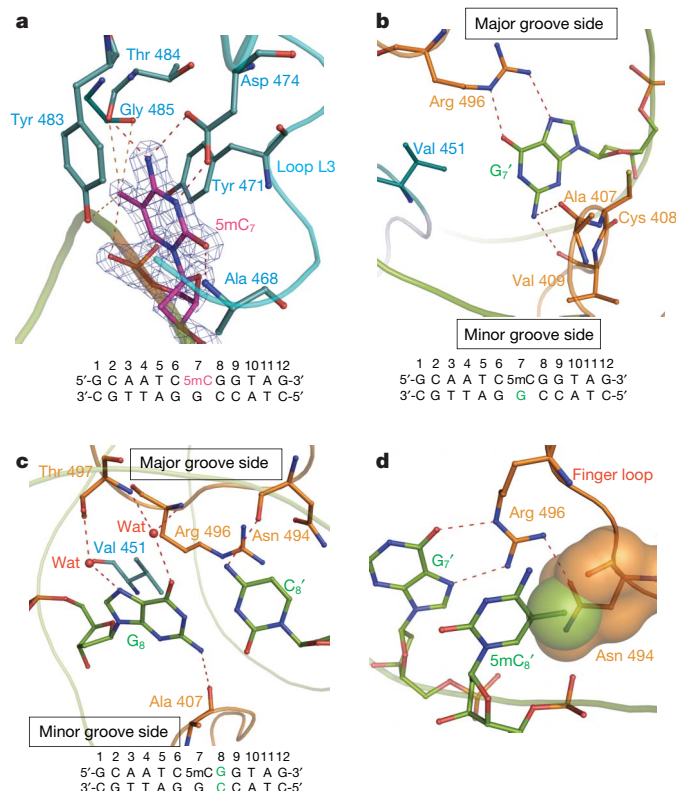


Figure 3 | Specific recognition of a hemi-methylated CpG site by SRA. The colour codes for the protein and DNA are same as in Fig. 2a. Red dotted lines show hydrogen-bond contacts within 4 Å. **a**, Close-up view of the flipped-out 5mC7 recognition site. The composite-omit electron density map for 5mC7 (more than 6.0 σ) is shown in blue. **b**, The orphaned guanine G7' base is retained near a canonical position in the DNA duplex by interactions with the SRA residues. **c**, Interactions between the G8•C8' base pair adjacent to the flipped-out base and the SRA residues. Wat, water. **d**, The model of fully methylated DNA bound to the SRA. The methyl group of 5mC8' (shown as a green sphere) causes steric interference with Asn 494 (orange spheres) in the finger loop. The hydrogen-bond distances between these bases and the SRA residues are listed in Supplementary Tables 2 and 3.

single methylcytosine base (250 Å²) at the hemi-methylated site. Thus, the base-flipping mechanism seems to allow SRA to gain sufficient interfacial area with only one methylcytosine. It may be that factors associated with the replication machinery facilitate base flipping from the helix. The resulting discrimination of methylation status by SRA is probably required for faithful epigenetic inheritance of the genome methylation status.

We next examined the coordinated actions of UHRF1 and the maintenance methyltransferase Dnmt1 on hemi-methylated DNA. The catalytic domain of Dnmt1 is thought to have a similar folding and mode of DNA interaction to those of Dnmt family members and

bacterial DNA methyltransferases⁸. Using the structure of the *HhaI* methyltransferase–DNA complex¹⁶ as a template, we constructed a model of the catalytic domain of Dnmt1 bound to DNA containing a hemi-methylated site. Dnmt1 methylates the cytosine adjacent to the 5mC•G pair in the hemi-methylated site. In the model, this cytosine swings out to fit into a pocket in the catalytic domain of Dnmt1. Overlaying the model on the crystal structure of the SRA–DNA complex suggests that the catalytic domain of Dnmt1 approaches the hemi-methylation site on the side opposite the SRA domain (Supplementary Fig. 7). The model predicts that the main bodies of the SRA and the Dnmt1 catalytic domains can coexist on DNA without a steric clash, but the base-recognition ('finger') loops of the proteins do face steric interference in the major groove. It therefore does not seem likely that SRA can remain in a specific complex with DNA when Dnmt1 approaches to exert its catalytic action on the target cytosine. This assumption is supported by the observation that SRA recognizes the target cytosine C8' when it is located within the helix, and thus the flipping out of this base by Dnmt1 presumably reduces the DNA-binding affinity of the SRA domain. It therefore seems unlikely that both the methylated and target cytosines simultaneously flip out when Dnmt1 binds to the preformed SRA–DNA complex. Rather, the methylcytosine and the target cytosine are more likely to swing out successively from the helix for recognition of methylation status and maintenance methylation, respectively, while the DNA is handed over from UHRF1 to Dnmt1 (Fig. 4). These concerted actions may ensure that methylation marks are accurately copied from the template to the newly synthesized daughter strands. The SRA domain has been shown to provide an interface with Dnmt1 (ref. 19); hence the domain has dual functions in DNA recognition and the recruitment of Dnmt1 to the proper position on the DNA.

METHODS SUMMARY

Crystallization and structure determination of the unliganded SRA domain.

The unliganded SRA domain (5.2 mg ml⁻¹), containing amino-acid residues 404–613 (Ser 404 is derived from the vector sequence), was crystallized at 4 °C by the hanging-drop vapour-diffusion method with 12% PEG3350 as a precipitant in 0.1 M Tris-HCl buffer pH 8.0 containing 0.1 M sodium acetate. The crystal was flash-frozen at 100 K in cryoprotectant containing 20% ethylene glycol. X-ray diffraction data were collected at a wavelength of 1.0000 Å on beamline BL-5A at Photon Factory, Tsukuba, Japan. Diffraction data were processed with the program HKL2000 (ref. 20). The structure of the SRA was solved by a molecular replacement method with the program molrep from the CCP4 suite²¹, using the crystal structure of the human UHRF1 SRA domain (PDB code 3BI7) as a search model. The initial model was built with the program O²², and refinement was performed against diffraction data at 1.77 Å with the programs CNS²³ and REFMAC²⁴. Detailed data collection and refinement statistics are given in Supplementary Table 1.

Crystallization and structure determination of the SRA–DNA complexes.

Co-crystals with DNA were obtained with two different 12-bp oligonucleotide duplexes that contained a single hemi-methylated CpG site (designated CpG-1 and CpG-2 in Supplementary Fig. 1a). Each complex structure was determined by molecular replacement, using the unliganded structure of the SRA as a search model. The structures of the monoclinic and tetragonal forms of the SRA bound

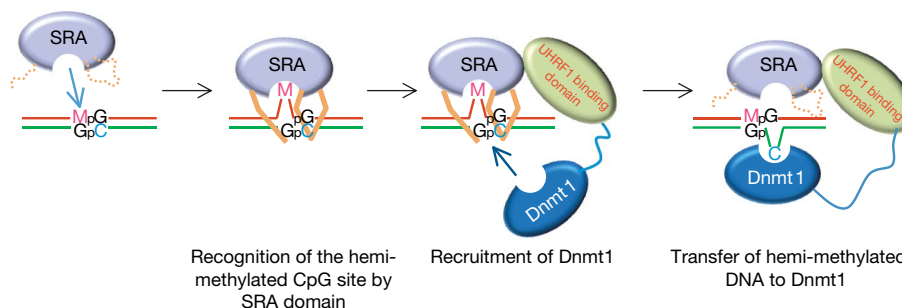


Figure 4 | A successive DNA transfer model for maintenance DNA methylation by UHRF1 and Dnmt1. Schematic model showing cooperative action by UHRF1 and DNMT1 for maintenance methylation (see text).

Pre-existing and newly synthesized DNA strands are indicated by red and green lines, respectively.

to hemi-methyl CpG-1 DNA were determined at 1.6 and 2.6 Å resolutions, respectively. The complex structure with hemi-methyl CpG-2 was determined at 2.55 Å resolution.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 March; accepted 9 July 2008.

Published online 3 September 2008.

- Chen, R. Z. *et al.* DNA hypomethylation leads to elevated mutation rates. *Nature* **395**, 89–93 (1998).
- Bird, A. & Wolffe, A. P. Methylation-induced repression—belts, braces, and chromatin. *Cell* **99**, 451–454 (1999).
- Ballabio, A. & Willard, H. F. Mammalian X-chromosome inactivation and the XIST gene. *Curr. Opin. Genet. Dev.* **2**, 439–447 (1992).
- Surani, M. A. Imprinting and the initiation of gene silencing in the germ line. *Cell* **93**, 309–312 (1998).
- Xu, G. L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402**, 187–191 (1999).
- Costello, J. F. *et al.* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nature Genet.* **24**, 132–138 (2000).
- Tajima, S. & Suetake, I. Regulation and function of DNA methylation in vertebrates. *J. Biochem.* **123**, 993–999 (1998).
- Hermann, A., Gowher, H. & Jeltsch, A. Biochemistry and biology of mammalian DNA methyltransferases. *Cell. Mol. Life Sci.* **61**, 2571–2587 (2004).
- Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
- Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
- Ooi, S. K. & Bestor, T. H. Cytosine methylation: remaining faithful. *Curr. Biol.* **18**, R174–R176 (2008).
- Ohki, I. *et al.* Solution structure of the methyl-CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* **105**, 487–497 (2001).
- Ho, K. L. *et al.* MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* **29**, 525–531 (2008).
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
- Suck, D. DNA–protein interactions. Flip out and modify. *Curr. Biol.* **4**, 252–255 (1994).
- Cheng, X. & Blumenthal, R. M. Finding a basis for flipping bases. *Structure* **4**, 639–645 (1996).
- O’Gara, M., Roberts, R. J. & Cheng, X. A structural basis for the preferential binding of hemimethylated DNA by HhaI DNA methyltransferase. *J. Mol. Biol.* **263**, 597–606 (1996).
- Tubbs, J. L., Pegg, A. E. & Tainer, J. A. DNA binding, nucleotide flipping, and the helix–turn–helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst.)* **6**, 1100–1115 (2007).
- Achour, M. *et al.* The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27**, 2187–2197 (2008).
- Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
- Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Morikawa for discussion, and N. Matsugaki, N. Igarashi, Y. Yamada, M. Suzuki and S. Wakatsuki for data collection at PF-BL5. This work was supported by grants to M.S. from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan and the Japan Science and Technology Agency, and by a Grant-in-Aid for Scientific Research to K.A. from the Japan Society for the Promotion of Science. The authors acknowledge support from the Global COE Program ‘Integrated Materials Science’ of MEXT of Japan.

Author Information Atomic coordinates and structural factors for the reported structures have been deposited with the Protein Data Bank under accession codes 2ZKG (unliganded SRA), 2ZKD (SRA-5mCpG-1 DNA, monoclinic form), 2ZKE (SRA-5mCpG-1 DNA, tetragonal form), 2ZKF (SRA-5mCpG-2 DNA). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.S. (shirakawa@moleng.kyoto-u.ac.jp) or M.A. (ma4@eng.mbox.media.kyoto-u.ac.jp).

METHODS

Protein expression and purification. DNA fragments encoding the SRA domain of mouse UHRF1 were amplified by PCR and cloned into the *Escherichia coli* expression vector pGEX4T-3 (GE Healthcare Bio-Sciences), which has been engineered for protein expression with the N-terminal glutathione *S*-transferase (GST) and small ubiquitin-like modifier-1 (SUMO-1) fusion tag. The SRA domain containing amino-acid residues 392–613 was expressed and purified as described below, but the N-terminal region (residues 392–404) was susceptible to proteolysis. The truncated SRA domain consisting of residues 405–613 was therefore reconstructed and overexpressed in *E. coli* strain BL21 (DE3) as a GST–SUMO-1 fusion protein. Cells were grown at 37 °C and induced with 0.2 mM isopropyl β -D-thiogalactoside (IPTG) to an attenuation of 0.5–0.6 at 660 nm. After incubation at 15 °C for 15 h, the cells were harvested, resuspended in lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT), 10% glycerol, 0.2 mM phenylmethylsulphonyl fluoride (PMSF)) and disrupted by sonication at 4 °C. After centrifugation, the supernatant was applied to a GST affinity column of Glutathione Sepharose 4 fast flow (GE Healthcare Bio-Sciences) equilibrated with the PMSF-free lysis buffer, and the GST–SUMO-1 fusion protein was eluted with elution buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM DTT, 10% glycerol, 30 mM reduced glutathione). The GST–SUMO tag was removed by digestion with the SUMO-specific protease GST–SEN2 for 12 h at 4 °C. The protein was further purified by cation-exchange and affinity column chromatography with HiTrap SP HP (GE Healthcare Bio-Sciences) and HiTrap Heparin HP (GE Healthcare Bio-Sciences), respectively. Finally, the protein solution was applied to a size-exclusion Hiload 16/60 Superdex 75 column (GE Healthcare Bio-Sciences) equilibrated with 10 mM HEPES pH 7.0, 150 mM NaCl, 1 mM DTT. Purified protein was concentrated at 4 °C to 10 mg ml⁻¹ in an Amicon Ultra 5,000 cut-off membrane concentrator (Millipore).

DNA binding assay. Each synthesized oligonucleotide (BEX Co. Ltd) was dissolved in 10 mM HEPES buffer pH 7.0 to a final concentration of 2 mM, and then the mixtures of equimolar complementary oligonucleotides were heated at 95 °C for 5 min and annealed at 4 °C for 12 h. Samples (10 μ l) containing 40 pmol of the SRA domain and 20 pmol of each DNA were incubated for 30 min at 4 °C in binding buffer (10 mM HEPES pH 7.0, 150 mM NaCl, 1 mM DTT, 1 mM EDTA, 10% glycerol), and then analysed on a native 8% polyacrylamide gel. Double-stranded co-polymer poly(dI•dC)•poly(dI•dC) (GE Healthcare Bio-Sciences) (120 ng) was used as a nonspecific DNA competitor. Electrophoresis was performed in running buffer (50 mM Tris-HCl pH 7.0, 45 mM boric acid, 0.5 mM EDTA) at 10 mA for 30 min in a cold room. DNA bands were detected by staining with GelRed (New England Biolabs).

Crystallization, data collection, structure determination, and refinement of the SRA–DNA complex. The screening trials of co-crystallization were performed with a 160 μ M concentration of the SRA–DNA complex. The monoclinic crystals of the SRA–hemi-methyl CpG-1 complex were obtained with a reservoir solution containing 0.1 M sodium citrate (pH 5.6), 0.2 M sodium acetate and 30% PEG4000. The tetragonal crystals were grown from a reservoir solution containing 0.1 M Bis-Tris propane pH 7.5, 0.2 M sodium fluoride and 20%

PEG3350. The SRA–hemi-methyl CpG-2 complex was crystallized in space group *P*₄₁₂₁₂ at 20 °C against a reservoir containing 0.1 M Bis-Tris propane pH 7.5, 0.2 M sodium fluoride and 20% PEG3350. The monoclinic *C*2 crystals contain two complexes in an asymmetric unit, whereas the tetragonal *P*₄₁₂₁₂ crystals have one complex. All crystals were flash-frozen at 100 K in cryoprotectant containing 20% ethylene glycol, and X-ray diffraction data sets were collected at a wavelength of 1.0000 Å on beamline BL-5A at Photon Factory. The crystals of the SRA–hemi-methyl CpG-1 DNA complex in space groups *C*2 and *P*₄₁₂₁₂ and those of the SRA domain–hemi-methyl CpG-2 DNA complex diffracted up to 1.60, 2.60 and 2.55 Å resolution, respectively. Each diffraction data set was processed with the program HKL2000. The crystallographic data and data collection statistics are given in Supplementary Table 1.

The structure of the SRA–hemi-methyl CpG-1 DNA complex in monoclinic *C*2 form was solved by a molecular replacement method with the unliganded SRA structure as a search model. The initial model was built with the program O and was refined against diffraction data at 1.6 Å with simulated annealing and energy minimization methods by using CNS (Supplementary Fig. 8). The final model contains 404–613 amino-acid residues of the SRA domain, in which Ser 404 is derived from the vector sequence, and 12-bp oligonucleotides. Two other structures of the SRA–hemi-methyl CpG complexes in a tetragonal form were determined with a similar procedure. The final model of the SRA–hemi-methyl CpG-2 DNA includes the amino-acid residues 405–613 and the 12-bp oligonucleotide except for thymines at 5' in both strands, T₁ and T₁₂'. Refinement statistics are given in Supplementary Table 1.

All figures and surface potentials of the protein were generated by PyMol²⁵. Accessible contact areas between the SRA domain and DNA were calculated from the protein–protein interaction server²⁶. The protein–DNA contacts were analysed with the program nucplot²⁷.

Model building of the SRA–DNA–Dnmt1 ternary complex. The comparative model of the DNMT1 catalytic domain (residues 1141–1619) was generated with the program phyre, version 0.2 (<http://www.sbg.bio.ic.ac.uk/~phyre/>), with the crystal structure of bacterial methyltransferase *HhaI* in the DNA-bound form (PDB code 1MHT). The primary sequence similarity between the catalytic domains of DNMT1 and *HhaI* is 42.8% (17.5% identity), whereas the sequence similarity of the catalytic domains between the DNMT family members is in the range 35–48% (13–16% identity). The DNMT1–DNA complex model was generated by superimposing the comparative model of DNMT1 on *HhaI* in complex with hemi-methylated DNA. Subsequently, the model of the SRA–hemi-methylated DNA–Dnmt1 ternary complex was built by superimposing the DNA phosphate backbone atoms around the target cytosine in the Dnmt1–DNA complex model on the corresponding atoms in the SRA–DNA complex, namely the backbone atoms of A₄–T₁₀ and G₇'–C₁₂' in the SRA–DNA complex.

25. DeLano, W. L. *The PyMOL Molecular Graphics System* (DeLano Scientific, 2002).

26. Jones, S. & Thornton, J. M. Protein–protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.* **63**, 31–59 (1995).

27. Luscombe, N. M., Laskowski, R. A. & Thornton, J. M. NUCPLOT: a program to generate schematic diagrams of protein–nucleic acid interactions. *Nucleic Acids Res.* **25**, 4940–4945 (1997).

LETTERS

Structural basis for recognition of hemi-methylated DNA by the SRA domain of human UHRF1

George V. Avvakumov^{1*}, John R. Walker^{1*}, Sheng Xue¹, Yanjun Li¹, Shili Duan², Christian Bronner³, Cheryl H. Arrowsmith^{1,2} & Sirano Dhe-Paganon^{1,4}

Epigenetic inheritance in mammals is characterized by high-fidelity replication of CpG methylation patterns during development^{1,2}. UHRF1 (also known as ICBP90 in humans and Np95 in mouse)³ is an E3 ligase important for the maintenance of global and local DNA methylation *in vivo*^{4,5}. The preferential affinity of UHRF1 for hemi-methylated DNA over symmetrically methylated DNA by means of its SET and RING-associated (SRA) domain⁶ and its association with the maintenance DNA methyltransferase 1 (DNMT1) suggests a role in replication of the epigenetic code^{4,5,7}. Here we report the 1.7 Å crystal structure of the apo SRA domain of human UHRF1 and a 2.2 Å structure of its complex with hemi-methylated DNA, revealing a previously unknown reading mechanism for methylated CpG sites (mCpG). The SRA–DNA complex has several notable structural features including a binding pocket that accommodates the 5-methylcytosine that is flipped out of the duplex DNA. Two specialized loops reach through the resulting gap in the DNA from both the major and the minor grooves to read the other three bases of the CpG duplex. The major groove loop confers both specificity for the CpG dinucleotide and discrimination against methylation of deoxycytidine of the complementary strand. The structure, along with mutagenesis data, suggests how UHRF1 acts as a key factor for DNMT1 maintenance methylation through recognition of a fundamental unit of epigenetic inheritance, mCpG.

Although DNMT1 shows a preference for hemi-methylated versus symmetrically methylated DNA substrates of up to 30–40-fold *in vitro*⁸, it has been argued that this is insufficient to account for the observed high fidelity of the inheritance of DNA methylation patterns⁹. To ensure this, it is thought that UHRF1 facilitates the loading of DNMT1 onto appropriate sites of newly replicated DNA⁹. To gain insight into how the SRA domain of UHRF1 recognizes methylated CpG sequences, we performed a crystallographic study of this domain.

The structure of the SRA domain of human UHRF1 was determined in its free form (Supplementary Fig. 1) and in complex with a 12-base-pair (bp) double-stranded DNA with a central hemi-methylated CpG (Fig. 1a) and refined to 1.7 Å and 2.2 Å resolution, respectively (Supplementary Table 1). The SRA domain appears to be a new DNA-binding fold, consisting of a 210-residue twisted, flared β-barrel. The sides of the barrel are covered with α-helical elements: α2 at one side connects strands β5 and β6, and a helix-loop pair at the other side connects strands β6 and β7 (Fig. 1a and Supplementary Fig. 1a). The extensive twist of the barrel partially separates strands β2 and β4 and forms the centre of the interaction surface with the hemi-methylated DNA, which runs parallel to the β-barrel along the β2 strand. The overall DNA-interaction mode can be thought of as a hand grasping

at the duplex DNA, with a methylcytosine-binding pocket in the palm formed by the distorted strands β2 and β4, and two

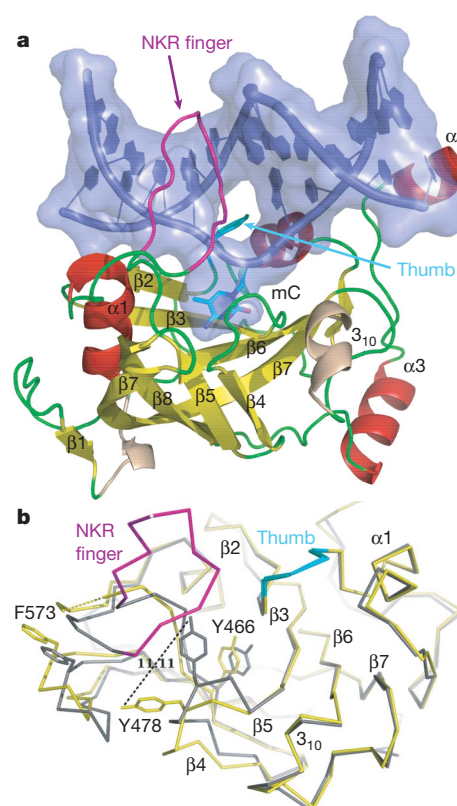


Figure 1 | Structure of the SRA domain of UHRF1 and its DNA-bound form.

a, The DNA–SRA complex. The SRA domain is shown and labelled with flexible loops, β-strands and α-helices in green, yellow and red, respectively. The DNA duplex is shown as a surface representation with backbone atoms and bases coloured blue. The flipped, methylated deoxycytidine (C6) is labelled as ‘mC’; the resulting hole in the DNA is partially filled by the NKR finger (magenta) and the thumb (α1–β2 loop, cyan). **b**, Conformational changes on DNA binding. Superposition of the Cα-only ribbon structure of apo (yellow) and DNA-bound (grey) SRA, highlighting conformational and dynamic changes on DNA binding, particularly in the methylcytosine-binding pocket, where the β5–α2 loop becomes ordered and the β3–β4 loop clamps down on the methylated cytosine. The black dashed line indicates the movement of the Tyr 466 hydroxyl towards the DNA in the DNA-bound structure relative to the apo SRA structure.

¹Structural Genomics Consortium, University of Toronto, 100 College Street, Toronto, Ontario M5G 1L5, Canada. ²Ontario Cancer Institute and Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, Ontario M5G 1L7, Canada. ³CNRS UMR 7175, IGL, Université Louis Pasteur Strasbourg I, Département de Pharmacologie et Pharmacochimie des Interactions cellulaires et moléculaires, Faculté de Pharmacie, 74 route du Rhin, B.P. 60024, 67401 Illkirch, France. ⁴Department of Physiology, University of Toronto, 100 College Street, Toronto, Ontario M5G 1L5, Canada.

*These authors contributed equally to this work.

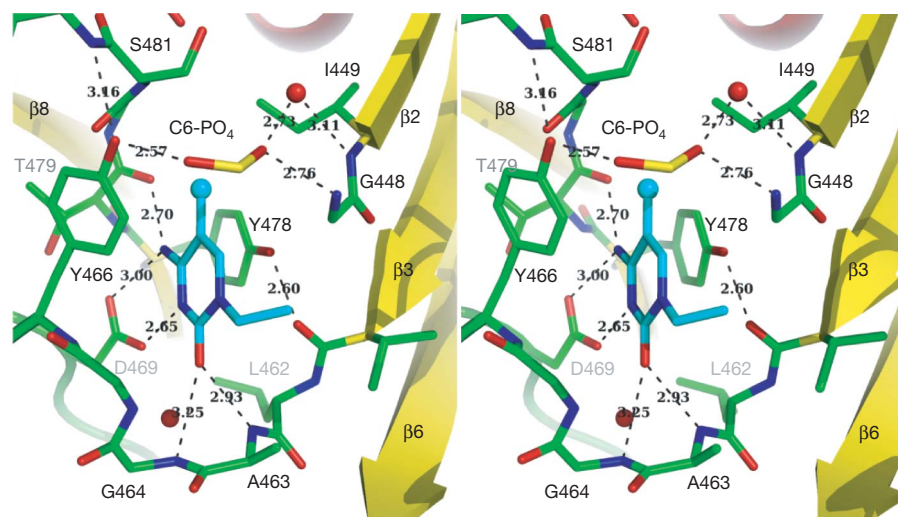


Figure 2 | The methylcytosine-binding pocket.

A three-dimensional, close-up stereo image of the methylcytosine (mC6, coloured in cyan) in the binding pocket, showing interacting residues in stick format. Also included is the phosphate group of mC6 and two well-defined water molecules (red spheres) in the pocket. The 5-methyl group is shown as a small sphere. Hydrogen bonds are shown as dashed lines and labelled with distances (Å). Nearby secondary structure elements are drawn and labelled. The deoxyribose of mC6 and the remaining double-stranded DNA are not shown for clarity.

loops corresponding to a finger and a thumb; the former projects into the major groove and the latter into the minor groove of the double helix.

The specialized binding pocket of the SRA domain is tailored for recognition of 5-methylcytosine, which is flipped out of the duplex DNA. It is about $6 \times 10 \times 6$ Å (width \times length \times depth), the walls being formed by the region between strands $\beta 2$ and $\beta 4$ and the $\beta 3$ – $\beta 4$ loop. The methylcytosine-binding pocket has four features that allow for selective binding (Fig. 2 and Supplementary Fig. 2). First, it is too small to fit purines. Second, two aromatic residues that line the pocket, Tyr 478 and Tyr 466, sandwich the methylcytosine by π -stacking interactions. Third, Asp 469, situated at the bottom of the pocket, makes two hydrogen bonds with N3 and N4 of the methylcytosine, which, in combination with a hydrogen bond between N4 and the main-chain carbonyl of Thr 479, confers preferential recognition of cytosine over thymine. The methylcytosine carbonyl makes hydrogen bonds with main-chain nitrogens of Ala 463 and Gly 464. Finally, specificity for methylcytosine over cytosine is conferred by a hemisphere of approximately 2 Å radius, which exactly fits the methyl group (Supplementary Fig. 3 and Supplementary Table 2). This methyl-group-binding site is deep within the methylcytosine pocket and would be occupied by solvent and therefore unfavourably hydrated if non-methylated cytosine were the ligand.

This precisely organized methylcytosine pocket has a significantly smaller volume compared to that present in the unbound form of the protein, owing to conformational changes of residues 464–473, 567–576 and 478–501, including the movement of the Tyr 466 side chain by 11 Å (measured from OH to OH; Fig. 1b and Supplementary Fig. 4). The presence of a glycine at position 448 at the entrance to the pocket seems to be crucial, because any other residue in this position would probably block entry to the pocket. Indeed, mutation of Gly 448 to aspartic acid resulted in a protein with significantly lower affinity for DNA (Fig. 3a). Other individual mutations in the pocket that reduced side chains to alanine did not significantly affect affinity, hinting at a complex network of cooperative interactions that involve multiple residues contributing to ligand-binding affinity and specificity (Fig. 3a). This methylcytosine recognition mode is in marked contrast to that of the methyl-binding domain of MeCP2, which recognizes mCpG without base flipping by means of the distinct hydration pattern of the symmetrically di-methylated mCpG in the major groove of duplex DNA¹⁰.

The finger and the thumb of the SRA domain also have key roles in recognition of the DNA duplex. The $\beta 5$ – $\alpha 2$ loop (referred to here as the NKR finger because of the Asn-Lys-Arg sequence located at its tip) enters the major groove and forms extensive hydrogen bonds with the components of the CpG sequence (Figs 1a and 4). The minor groove is

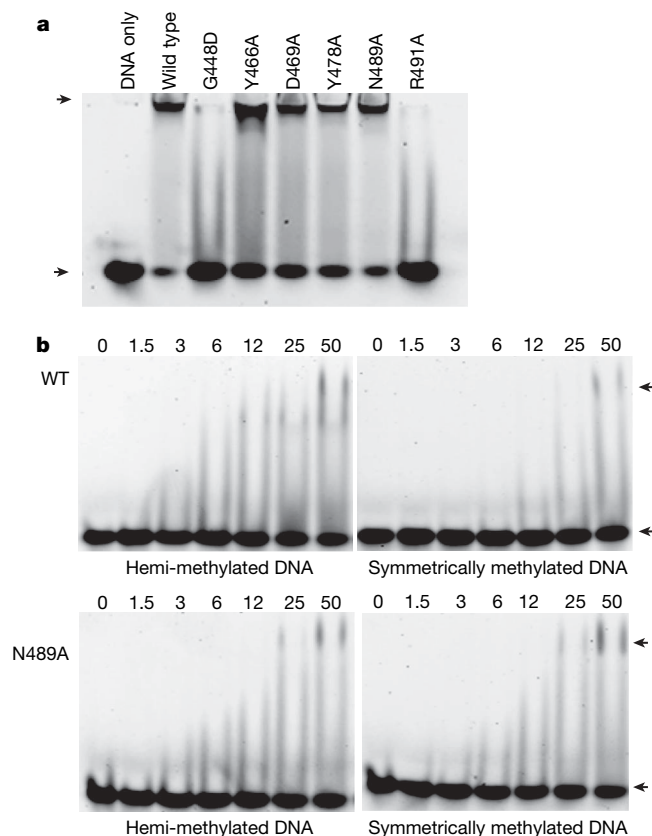


Figure 3 | SRA domain specificity determinants. **a**, Electrophoretic mobility shift assay of SRA domain with hemi-methylated DNA. Human UHRF1 fragment 414–617, either wild type or mutated as indicated, was incubated with fluorescently-tagged, hemi-methylated, 12-bp duplex DNA and separated on a polyacrylamide gel. Positions of free DNA (lower bands) and protein–DNA complexes (upper bands) are indicated with arrows. Equal amounts of purified proteins (100 pmol per lane) and DNA (1 pmol per lane) were used in each lane. **b**, N489A mutation eliminates ligand specificity. Increasing concentrations of UHRF1 fragment 121–617 (wild type, WT) and its N489A mutant were incubated with hemi-methylated or symmetrically di-methylated, fluorescently-tagged, 12-bp DNA, and separated as above. Each lane contained 1 pmol DNA, and molar excess of protein over the DNA is indicated above the lane.

occupied by the helix–loop pair that connects strands $\beta 1$ and $\beta 2$ of the barrel (the thumb; Supplementary Fig. 5). This segment interacts in a sequence-independent manner with the backbone atoms of the DNA; the thumb has no specific interactions with the DNA bases. Remarkably, the thumb interacts with the finger through the gap in the double-stranded DNA left by the extrusion of the 5-methyl-deoxycytidine. A methyl group of Val 446 in the thumb reaches through to form a van der Waals interaction with the aliphatic portion of the Arg 491 side chain in the NKR finger on the opposite side of the duplex DNA. Thus, the thumb may contribute to expulsion of the 5-methyl-deoxycytidine (Fig. 1a and Supplementary Fig. 5). Surprisingly, the gross features of the SRA–DNA interaction are reminiscent of DNA methyltransferases and other DNA-modifying enzymes¹¹ as opposed to sequence-specific DNA-binding factors that recognize sequence-dependent features of duplex DNA. For example, glycosidases¹², methyltransferases^{13,14} and other DNA- or RNA-modifying enzymes¹⁵ require base flipping to access the target base for catalysis.

CpG sequences are unique in the mammalian genome in that their methylation status carries critical epigenetic information. In the complex structure, the NKR finger interrogates the major groove surrounding the flipped methyldeoxycytidine, recognizing both the gap in the DNA and its context within the duplex CpG sequence (Figs 1a and 4b). Notably, Arg 491 in the finger substitutes for methyldeoxycytidine, and hydrogen bonds to the orphaned guanine (G6'; see Fig. 4a for labelling of DNA bases); the guanidinium side chain forms two hydrogen bonds with the nitrogen atom at position 7 (N7) and oxygen atom at position 6 (O6) of the orphaned guanine. Moreover, Arg 491 stacks with the bases C5 and G7 on the methylated strand, thereby compensating for the extrusion of the methyldeoxycytidine from the DNA double helix and ensuring specificity for a methylated cytosine (Fig. 4). Furthermore, the NKR finger interacts with the G7•C7' base pair through a combination of hydrogen bonds and favourable van der Waals interactions (Fig. 4). These interactions occur in the context of a normal duplex DNA because the G7•C7' base pair is engaged in characteristic hydrogen

bonds (although the bases are slightly out-of-plane relative to each other). Recognition of the duplex CpG sequence is further ensured by the ability of Arg 491 to simultaneously recognize G6' (through guanidinium side-chain interactions) and G7 (through peptide backbone interactions). Thus, Arg 491 clearly has a central role in both affinity and specificity for the 5-mCpG sequence. This was confirmed by mutation of Arg 491 to alanine resulting in a marked decrease of binding to methylated DNA (Fig. 3a). The capture of the methylated base in the specificity pocket and its replacement in the base stack by Arg 491 is probably a method by which the SRA domain can make a tight interaction with its target in the absence of many base-specific contacts. Whether the NKR finger, and Arg 491 in particular, is actively involved in flipping the methylated base, or recognizes a gap in the DNA after flipping, requires further study.

Our structure also suggests that the NKR finger may be able to recognize the absence of a methyl group on C7' of the complementary strand in addition to recognizing the presence of this modification on the target strand (mC6). Indeed, the side chains of Asn 489 and Arg 491 not only recognize the orphaned guanine (G6') and the adjacent, non-methylated cytosine (C7') but they also work together to effectively mask the C5 position of the non-methylated C7' cytosine. The amide side chain of Asn 489 hydrogen bonds with the guanidinium group of Arg 491 as well as with the backbone phosphate group of C7', placing the C5 atom of the non-methylated cytosine within van der Waals distance (3.1 Å and 3.4 Å) of the backbone oxygen and amide side-chain nitrogen, respectively, of Asn 489 (Fig. 4b and Supplementary Fig. 5). This suggests that methylation at the C5 atom of C7' is likely to perturb the position of Asn 489 (or the base position), thereby disrupting the intricate network of interactions at the NKR finger–DNA interface and destabilizing the conformational state of Arg 491, which acts as a critical anchor in our structure. This is in agreement with recent data showing that the mouse UHRF1 SRA domain has a sevenfold greater affinity for multiply hemi-methylated double-stranded DNA compared with fully methylated double-stranded DNA⁵. We also observed that human

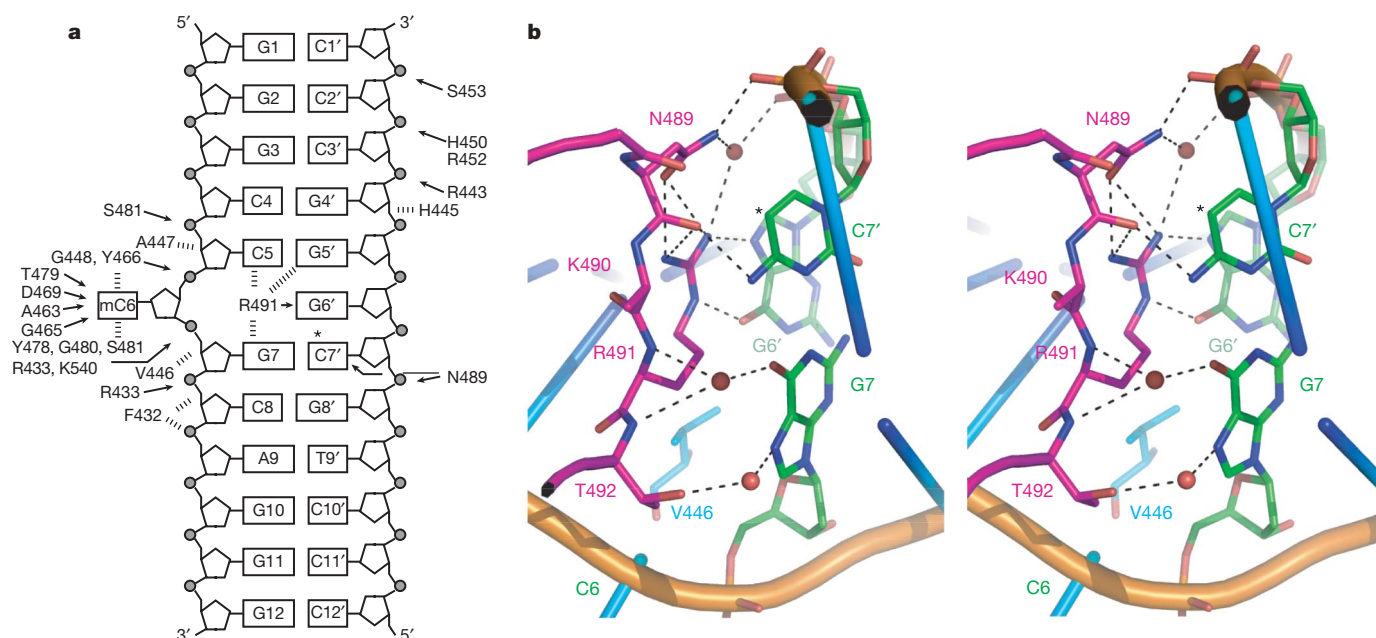


Figure 4 | Role of the NKR finger in the DNA binding. **a**, Schematic representation of the interactions with the DNA. Van der Waals contacts to DNA are shown as a series of short, parallel lines, and hydrogen bonds are shown as arrows. The carbon atom at position 5 of C7' is marked with an asterisk in this and subsequent panels. **b**, Close-up view of the finger–double-stranded DNA interface in three-dimensional stereo view. The NKR finger and thumb, coloured as in Fig. 1, and bases (green) are shown in stick format. The double-stranded DNA backbone is coloured

brown and adjacent bases are represented in PyMol cartoon format. Three water molecules shown as red spheres mediate interactions, two of which are between the finger and G7 and the other molecule coordinates the backbone of the DNA and the side chains of Asn 489 and Arg 491. The C5 carbon atom of the non-methylated C7' (marked with an asterisk) is buried by the NKR finger and is within van der Waals distances from side chain and backbone atoms of Asn 489. For clarity, the side chain of Lys 490, which is solvent exposed, is not shown.

UHRF1 constructs encompassing the SRA domain show a modest preference for binding to a 12-mer double-stranded DNA hemi-methylated at a single site, compared to a symmetrically di-methylated substrate (data not shown).

To assess the extent to which Asn 489 contributes to discrimination between hemi-methylated versus symmetrically methylated CpG sites, we examined the ability of an N489A mutant protein, encompassing UHRF1 residues 121–617, to bind to a hemi-methylated 12-mer double-stranded DNA substrate and the corresponding symmetrically di-methylated substrate. The N489A protein bound equally well to both forms of DNA (Fig. 3b), in agreement with our steric clash model. Thus, the NKR finger has a role in both recognition of the CpG sequence and discrimination against methylation of the complementary strand. Interestingly, the plant SRA-domain-containing proteins have divergent finger residues (Supplementary Fig. 6) and recognize the methylation status of the complementary strand within different sequence contexts¹⁶.

The crystal structures described here demonstrate that the SRA domain of human UHRF1 is uniquely adapted for specific recognition of one of the most important epigenetic marks present in human and animal cells—methylated CpG sequences. Given the known ability of UHRF1 to bind to DNMT1, our findings support the hypothesis that UHRF1 is a factor involved in increasing the efficacy of DNMT1 methylation of hemi-methylated DNA and thus the mitotic inheritance of the CpG methylation patterns⁹. Additional activities of UHRF1, such as binding to H3K9me3, a hallmark of mCpG-rich heterochromatin, may also serve to co-localize UHRF1 and DNMT1 to genomic sites that require maintenance methylation¹⁷.

METHODS SUMMARY

The In-Fusion system was used to clone constructs of human UHRF1, which were expressed in bacterial cells and purified by metal-affinity chromatography and gel filtration. Mutants were prepared by the QuikChange (Stratagene) site-directed mutagenesis method. Crystals of the SRA apo protein and SRA complexed with hemi-methylated, double-stranded DNA were obtained using the hanging-drop technique, cryoprotected and frozen by immersion into liquid nitrogen. Crystallographic data were collected at the APS synchrotron beamline 23-ID-B. The apo structure was solved using single anomalous dispersion at the selenium edge and refined to a final R_{free} of 19.7% (1.7 Å), and the SRA–DNA complex structure was determined by molecular replacement and refined to a final R_{free} of 23.4% (2.2 Å). Electrophoretic mobility shift assays were run on 10% acrylamide gel at pH 8 and screened using fluorescence of an oligonucleotide labelled with an Alexa 647 tag at its 5' end.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 March; accepted 18 July 2008.

Published online 3 September 2008.

- Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226–232 (1975).
- Razin, A. & Riggs, A. D. DNA methylation and gene function. *Science* **210**, 604–610 (1980).
- Unoki, M., Bronner, C. & Mousli, M. A concern regarding the current confusion with the human homolog of mouse Np95, ICBP90/UHRF1. *Radiat. Res.* **169**, 240–244 (2008).

- Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
- Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
- Unoki, M., Nishidate, T. & Nakamura, Y. ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. *Oncogene* **23**, 7601–7610 (2004).
- Achour, M. *et al.* The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27**, 2187–2197 (2008).
- Yoder, J. A., Soman, N. S., Verdine, G. L. & Bestor, T. H. DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. *J. Mol. Biol.* **270**, 385–395 (1997).
- Ooi, S. K. & Bestor, T. H. Cytosine methylation: remaining faithful. *Curr. Biol.* **18**, R174–R176 (2008).
- Ho, K. L. *et al.* MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* **29**, 525–531 (2008).
- Roberts, R. J. On base flipping. *Cell* **82**, 9–12 (1995).
- Huffman, J. L., Sundheim, O. & Tainer, J. A. DNA base damage recognition and removal: new twists and grooves. *Mutat. Res.* **577**, 55–76 (2005).
- Cheng, X. & Blumenthal, R. M. Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**, 341–350 (2008).
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
- Hoang, C. & Ferre-D'Amare, A. R. Cocystal structure of a tRNA^{Psi55} pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell* **107**, 929–939 (2001).
- Johnson, L. M. *et al.* The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **17**, 379–384 (2007).
- Karagianni, P., Amazit, L., Qin, J. & Wong, J. ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Mol. Cell Biol.* **28**, 705–717 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank F. Chedin, A. Edwards and R. Klose for discussions, and A. Dong for collecting data at the synchrotron. Diffraction data were measured at the Advanced Photon Source (Argonne, Illinois), which was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract no. DE-AC02-06CH11357. The Structural Genomics Consortium is a registered charity (number 1097737) that receives funds from the Canadian Institutes for Health Research, the Canadian Foundation for Innovation, Genome Canada through the Ontario Genomics Institute, GlaxoSmithKline, Karolinska Institutet, the Knut and Alice Wallenberg Foundation, the Ontario Innovation Trust, the Ontario Ministry for Research and Innovation, Merck & Co., Inc., the Novartis Research Foundation, the Swedish Agency for Innovation Systems, the Swedish Foundation for Strategic Research and the Wellcome Trust. This research was also supported by the Canadian Cancer Society (S.D. and C.H.A.).

Author Contributions S.D.-P., G.V.A., J.R.W., C.B. and C.H.A. developed the experimental design. S.D.-P. designed and S.D. and Y.L. prepared plasmids for bacterial expression of recombinant proteins. G.V.A. and S.X. performed purification, characterization and crystallization of proteins and EMSA. J.R.W. designed the oligonucleotides and conducted the structural analysis. S.D.-P., C.H.A., J.R.W. and G.V.A. interpreted results and wrote the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession codes 3BI7 (SRA apo protein) and 3CLZ (SRA–DNA complex). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.D.-P. (sirano.dhepaganon@utoronto.ca).

METHODS

Cloning, protein expression, purification and crystallization. The complementary DNA encoding full-length human UHRF1 (BC113875.OBS. MHS4426-98361361.pCR-BluntII(TOPO)) was obtained from Open Biosystems, and its regions corresponding to the SRA domain (residues 414–617) and PHD-SRA tandem (121–617) were cloned into bacterial expression vectors (pNIC-CH in case of SRA domain and modified pET28a in case of the tandem domain) using ligase-independent cloning. The proteins were expressed in *Escherichia coli* BL21 (DE3) grown in Terrific Broth (TB) in the presence of 50 $\mu\text{g ml}^{-1}$ of kanamycin at 37 °C to a D_{600} of 4.5–5. Protein expression was induced with 0.2 mM isopropyl-1-thio- β -galactopyranoside and the cell cultures continued overnight (16–18 h) at 15 °C. The cultures were then centrifuged (12,000g, 15 min) and cell pellets collected and stored at –80 °C until used for protein purification.

The cell pellet from a 2 l culture was resuspended in 50 ml lysis buffer consisting of 10 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5% glycerol, 2 mM imidazole, 1 mM β -mercaptoethanol and 0.1 μM phenylmethyl sulfonyl fluoride, and cells were destroyed using a Microfluidizer Processor M-110EH (Microfluidics) at 18,000 pounds per square inch peak pressure. After centrifugation at 40,000g for 30 min, the clarified cell lysate was applied on a 3-ml column packed with TALON metal affinity resin (Clontech). The column was consequently washed with 10 ml wash buffer A (10 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5% glycerol, 10 mM imidazole, 1 mM β -mercaptoethanol), 10 ml wash buffer B (wash buffer A supplemented with 0.05% Tween 20) and 10 ml wash buffer A again. The protein was eluted with 6 ml 10 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5% glycerol and 200 mM imidazole, and dithiothreitol was added to the eluate to a final concentration of 2 mM. The protein was further purified by gel filtration on a HighLoad 16/60 Superdex 200 column (GE Healthcare) equilibrated with 20 mM Tris-HCl, pH 8.0, 0.5 M NaCl, 5% glycerol and 2 mM dithiothreitol. Fractions containing the target protein (analysed by SDS–PAGE) were pooled and concentrated by ultrafiltration using an Amicon Ultra centrifugal filter with 10 kD molecular weight cutoff limit to a final concentration of 30–40 mg ml^{-1} . Liquid chromatography-mass spectrometry (LC-MS) analysis was used to verify the molecular mass of the protein, and its purity was assessed by SDS–PAGE. The concentrated purified proteins were stored on ice for 2–3 weeks and at –80 °C for longer periods. To obtain the selenomethionyl derivative of the UHRF1 SRA domain, the cells were grown in M9 medium supplemented with glycerol using a M9 SeMET High-Yield growth media kit package (MD045004-50L, Medicilon) according to the manufacturer's instructions, and the protein was purified as above.

Mutated SRA cDNAs were made by using QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene); mutations were confirmed by sequencing complete cDNAs. Mutated proteins were expressed in bacterial cells and purified as above; molecular weights of mutated proteins (determined by LC-MS) were in line with the expected amino acid substitutions.

Crystals of the UHRF1 SRA domain selenomethionyl derivative were grown at 18 °C using the hanging-drop method by mixing one volume of protein solution at 34 mg ml^{-1} with one volume of well solution consisting of 1.4 M ammonium sulphate, 0.1 M bis-Tris, pH 6.0, 0.2 M NaCl and 1 mM TCEP. The crystals were cryoprotected by immersion in the well solution mixed in 1:1 ratio with a water solution containing 20% (w/v) sucrose, 4% (w/v) glucose, 18% (v/v) glycerol and 18% (v/v) ethylene glycol, and were placed into liquid nitrogen. To crystallize the SRA–DNA complex, the protein at 10 mg ml^{-1} was mixed, in a molar ratio of 1:1.5, with a hemi-methylated double-stranded DNA obtained by annealing two oligonucleotides, 5'-GGGCCXGCAGGG (X = 5-methyldeoxycytosine) and 5'-CCCTGCGGGCCC (Oligos Etc., Inc.), and the mixture was incubated for 1 h on ice. Crystals of the complex were grown at 18 °C using the hanging-drop method

by mixing one volume of the complex solution with one volume of well solution, consisting of 12% PEG1500, 0.2 M NaCl, 0.1 M bis-Tris, pH 7, and 1 mM TCEP, and 0.4 volume of 30% xylitol. The crystals were cryoprotected and frozen as above.

Electrophoretic mobility-shift assays. Fluorescent-labelled oligonucleotide, 5'-AGCTAXGATTGG, conjugated at its 5' end with Alexa 647 tag was synthesized and purified by Operon (X = 5-methyldeoxycytosine). To obtain duplex DNA, it was annealed to either 5'-CCAATCGTAGCT (hemi-methylated) or 5'-CCAATXGTAGCT (symmetrically di-methylated). Duplex DNA (1 pmole per lane) was preincubated (30 min on ice) with an appropriate concentration of protein (as indicated in the figure legend) in 20 mM Tris-Cl, pH 7.5, 200 mM NaCl, 1 mM DTT and 5% glycerol and subjected to polyacrylamide gel electrophoresis (10% BioRad gels run in 0.5 \times Tris-borate-EDTA buffer at 100 V for 1 h on ice). Gels were twice washed with water and scanned using an excitation wavelength of 633 nm (Typhoon Imager, GE Healthcare).

Data collection, structure solution and refinement. X-ray diffraction data from crystals of the selenomethionone-derivatized SRA domain were collected at APS synchrotron beamline 23-ID-B at a wavelength of 0.97943, and the structure was solved using the technique of single-wavelength anomalous diffraction. After locating the selenium substructure with the program SOLVE¹⁸, the program RESOLVE¹⁹ was used to build an initial model. Restrained, maximum likelihood refinement was carried out with the CCP4 program REFMAC²⁰. Manual model building was carried out using the program Coot²¹. At the later stages, TLS (translation, libration and screw-rotation parameterization) and restrained refinement was carried out, with the initial TLS parameters obtained from the TLSMD webserver²². The final model has an *R* factor of 0.159 and an *R*_{free} of 0.197 for data from 33.0 Å to 1.7 Å. The structure contains 1,785 atoms, including 238 solvent atoms. Missing residues include the initial selenomethionine, residues 484–495, and five residues of the C-terminal His tag.

X-ray diffraction data from crystals of the SRA domain in complex with a 12-bp double-stranded DNA were collected at APS synchrotron beamline 23-ID-B at a wavelength of 0.96863, and the structure was solved by molecular replacement using the program Phaser with the apo-SRA structure and ideal B-form DNA as search models. Restrained, maximum likelihood refinement was carried out with the CCP4 program REFMAC. Manual model building was carried out using the program Coot. In the later stages, TLS and restrained refinement was carried out, with the initial TLS parameters obtained from the TLSMD. The final model has an *R* factor of 0.189 and an *R*_{free} of 0.234 for data from 38.6 Å to 2.2 Å. The structure contains 9,116 atoms, including 643 solvent atoms and 1,948 DNA atoms. Four molecules of the protein–DNA complex are found in the asymmetric unit. All atoms of the double-stranded DNA have been located in the experiment. Missing protein residues include the initial methionine for each of the four molecules, the next residue for chains B and D, and the next two residues for chains A and C. On the C terminus, both A and C chains are complete to the end of the SRA domain sequence, but chains B and D have additional residues corresponding to the cloning tag.

18. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
19. Terwilliger, T. C. Maximum-likelihood density modification using pattern recognition of structural motifs. *Acta Crystallogr. D* **57**, 1755–1762 (2001).
20. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
21. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
22. Painter, J. & Merritt, E. A. TLSMD web server for the generation of multi-group TLS models. *J. Appl. Crystallogr.* **39**, 109–111 (2006).

LETTERS

The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix

Hideharu Hashimoto¹, John R. Horton¹, Xing Zhang¹, Magnolia Bostick², Steven E. Jacobsen^{2,3} & Xiaodong Cheng¹

Maintenance methylation of hemimethylated CpG dinucleotides at DNA replication forks is the key to faithful mitotic inheritance of genomic methylation patterns. UHRF1 (ubiquitin-like, containing PHD and RING finger domains 1) is required for maintenance methylation by interacting with DNA nucleotide methyltransferase 1 (DNMT1), the maintenance methyltransferase, and with hemimethylated CpG, the substrate for DNMT1 (refs 1 and 2). Here we present the crystal structure of the SET and RING-associated (SRA) domain of mouse UHRF1 in complex with DNA containing a hemimethylated CpG site. The DNA is contacted in both the major and minor grooves by two loops that penetrate into the middle of the DNA helix. The 5-methylcytosine has flipped completely out of the DNA helix and is positioned in a binding pocket with planar stacking contacts, Watson–Crick polar hydrogen bonds and van der Waals interactions specific for 5-methylcytosine. Hence, UHRF1 contains a previously unknown DNA-binding module and is the first example of a non-enzymatic, sequence-specific DNA-binding protein domain to use the base flipping mechanism to interact with DNA.

In mammals, DNMTs belong to two structurally and functionally distinct families³. The DNMT3 family establishes the initial CpG methylation pattern *de novo*, whereas DNMT1 maintains this pattern during chromosome replication^{4,5} and repair⁶. As a maintenance methyltransferase, DNMT1 has a preference for hemimethylated sites (discussed in ref. 7). DNMT1 is necessary but not sufficient for proper maintenance methylation. The SRA-domain-containing protein UHRF1, also known as NP95 (nuclear protein of 95 kDa) in mouse and ICBP90 (inverted CCAAT binding protein of 90 kDa) in human⁸, targets DNMT1 to hemimethylated replication forks^{1,2,9}. The UHRF1–DNMT1 interaction probably makes DNMT1 more specific for hemimethylated DNA than can be explained by inherent DNMT1 binding preferences alone¹⁰.

Here we show, by means of X-ray crystallography, that the SRA domain of mouse UHRF1 (residues 419–628; Supplementary Fig. 1) binds to a hemimethylated CpG dinucleotide and increases its protein–DNA interface by flipping 5-methylcytosine (5mC) into a specific binding pocket that probably prevents the SRA domain from sliding off. A 12-base-pair oligonucleotide (plus a 5′-overhanging thymine) containing a single CpG site methylated on one strand was used for co-crystallization with the SRA domain (Fig. 1a). We crystallized the SRA–DNA complex in two different space groups (*P*₂₁₂₁ and *P*₄₁₂₁) and determined the structures to the resolutions of 2.19 Å and 1.96 Å, respectively (Supplementary Table 1). The protein components of the two structures are highly similar, with a root mean squared deviation of less than 0.3 Å when comparing 206 pairs of C α atoms. Here we will describe the structure of *P*₄₁₂₁ and will discuss the differences between the two. In addition, we determined the structure of a non-specific SRA–DNA complex to a lower resolution at 3.09 Å in the space group *P*₆₁₂₂ (Supplementary Table 1).

The SRA domain contains two twisted β -sheets packed together to form a crescent moon-like structure (Fig. 1b). The 17-residue-long strand β 8 (Asn 557–Gly 573) is part of and links together the two sheets. Four helices (α 1– α 4) and one α ₁₀ helix are scattered through the entire domain, with helix α 1 packed against the first β -sheet, helix α 2 sandwiched between the two β -sheets, and helices α 3 and α 4

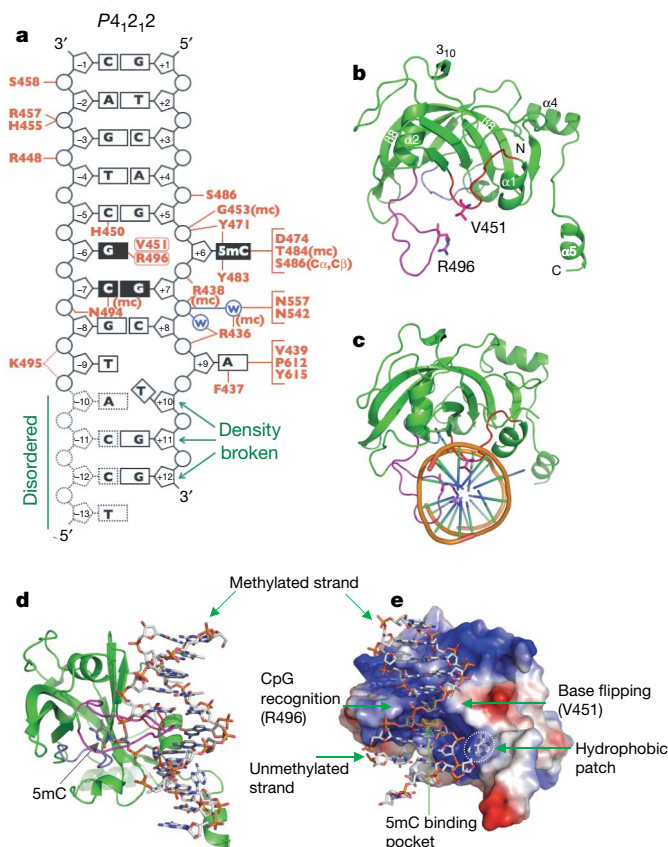


Figure 1 | Structure of SRA–DNA complex. **a**, Summary of the SRA–DNA interactions; mc, main-chain-atom-mediated contacts; w, water-mediated hydrogen bonds. Black boxes represent CpG recognition sequence and K495-associated dotted lines represent weak hydrogen bonds. **b**, The side chains of V451 of the base flipping loop and R496 of the CpG recognition loop are in direct van der Waals contact. **c**, The two loops—CpG recognition and base flipping—penetrate into the DNA helix from opposite directions. **d**, The 5mC flips out and binds in a cage-like pocket. **e**, The surface charge at neutral pH is displayed as blue for positive ($20 k_B T$), red for negative ($-20 k_B T$), and white for neutral, where k_B is the Boltzmann's constant and T is the temperature.

¹Department of Biochemistry, Emory University School of Medicine, 1510 Clifton Road, Atlanta, Georgia 30322, USA. ²Department of Molecular Cell and Developmental Biology, ³Hughes Medical Institute, University of California, Los Angeles, 621 Charles E. Young Dr. South, Los Angeles, California 90095, USA.

positioned on the inner and outer surfaces of the crescent, respectively. The carboxyl helix $\alpha 5$ protrudes away from the rest of the domain. The inner surface of the crescent contains three loops, one between helix $\alpha 1$ and strand $\beta 2$ (red), one between strands $\beta 3$ and $\beta 4$ (light blue), and one between strand $\beta 5$ and helix $\alpha 3$ (magenta). These loops are responsible for CpG recognition (Asn 494 and Arg 496 in magenta), promotion of base flipping (His 450 and Val 451 in red), and formation of a 5mC-binding pocket (Tyr 483–Ser 486 and Tyr 471–Asp 474; see later).

The DNA is bound to the basic inner surface of the crescent (Fig. 1c–e), with phosphate contacts spanning 8 base pairs but mostly concentrated on the 5 phosphates surrounding the 5mC (two 5' and three 3'; Fig. 1a). The SRA residues involved in phosphate interactions are scattered throughout an amino-terminal region of ~50 residues (436–486) by forming secondary structures of both β -sheets and their associated loops (Supplementary Fig. 2). The tips of two hairpin loops, Asn 494–Arg 496 and His 450–Val 451, approach the DNA from opposite directions. They contact major and minor grooves, respectively, and penetrate into the DNA helix by forming van der Waals contacts between side chains of Arg 496 and Val 451 in the centre of the DNA helix (Fig. 1c–e).

The SRA side chain intercalation is associated with 5mC flipping out of the DNA helix (Fig. 1d). Val 451 from the minor groove side occupies the space left by the everted base, and we refer to the Val 451-containing loop as the base-flipping-promotion loop (Fig. 1b–e). On the major groove side, Arg 496 makes bifurcated hydrogen bonding interactions with the intrahelical orphaned guanine (Fig. 2a). Asn 494 makes bridging hydrogen bonds between Arg 496 and the 5' phosphate of unmethylated cytosine on the opposite strand (Fig. 2a). Furthermore, the backbone carbonyl oxygen of Asn 494 is in close proximity to the ring carbon atom C5 of the unmethylated cytosine,

forming a $C=O \cdots H-C$ type of hydrogen bond. We name the Asn 494/Arg 496-containing loop the CpG recognition loop, as it seems to determine specificity for hemimethylated CpG. The addition of a methyl group to the ring carbon C5 of the unmethylated cytosine would cause a steric clash between the methyl group and Asn 494, and could account for the approximately sevenfold decrease in binding to fully methylated DNA¹ (Supplementary Fig. 3).

The extrahelical 5mC is bound in a hydrophobic cage formed by two tyrosines (Tyr 471 and Tyr 483) stacking the flipped base in between (Fig. 2b). The polar groups of the 5mC ring that normally form the Watson–Crick pair with guanine are now involved in hydrogen bonds with the main-chain amide nitrogen atoms of Ala 468 and Gly 470 (interacting with O2 oxygen atom), the side-chain carboxylate oxygen atoms of Asp 474 (interacting with N3 and N4 nitrogen atoms) and the main-chain carbonyl oxygen of Thr 484 (interacting with N4 nitrogen atom; Fig. 2c). Interactions with the exocyclic amino group N4 (NH_2) define the binding pocket base specificity for a cytosine. The distance between Asp 474 and the ring nitrogen N3 (2.7 Å) strongly suggests the presence of a hydrogen bond and therefore the presence of a proton either at the carboxylate oxygen or at N3. The most likely source of that proton is an ordered water molecule, which is in fact found in direct contact with the side chain of Asp 474 and a network of other water molecules. The methyl group at the C5 position in 5mC is in van der Waals contact with the $C\alpha$ and $C\beta$ atoms of Ser 486, the hydroxyl oxygen of which interacts with the phosphate of the 5' nucleotide (Fig. 2c).

Outside of the hemimethylated CpG dinucleotides, two additional bases are in direct contact with the SRA domain. The cytosine of the G•C pair at position 5 has a hydrogen bond, by means of its hydrogen-acceptor O2 in the minor groove, with His 450 (Fig. 2d). The His 450-mediated interaction is probably not base specific at this position, as in theory the hydrogen bond could also form with thymine O2 of an A•T pair, guanine N3 of a C•G pair, or adenine N3 of a T•A pair.

Surprisingly, the adenine at the position 9, two bases 3' to the 5mC, also adopts an extrahelical conformation (Fig. 1a) and stacks against a hydrophobic surface patch (Fig. 1e and Supplementary Fig. 4). The flipping of adenine seems to destabilize the rest of the DNA, leaving the last four nucleotides of the unmethylated strand (positions –10 to –13) disordered. In addition, the last two nucleotides of the methylated strand (positions 11–12)—having only broken, residual densities—appear as if they are single-stranded DNA. The flexibility of the last four base pairs (positions 9–12) is probably due to the fact that the SRA-domain–DNA interactions are shifted towards one end of the DNA, leaving the disordered end unattended (Fig. 1a). In the space group $P2_12_12_1$ crystal, these flexible base pairs (positions 9–12) were more ordered, stabilized by contacts with the crystallographic symmetry-related molecule (Fig. 3a and Supplementary Fig. 5). Interestingly, in this complex the adenine 9 adopts two conformations—an intrahelical one with normal pairing and the extrahelical one described for the $P4_12_12$ complex, with adenine 9 stacking against the hydrophobic patch. It is also possible that the hydrophobic patch normally interacts with residues outside of the SRA domain used here for crystallization¹¹. The absence of such residues and the weak crystallographic restraining forces on the end of the DNA may have caused adenine 9 to be extrahelical and interact with the hydrophobic patch in the crystal.

In the complex with unmethylated DNA, the CpG recognition loop (residues 490–500) appears disordered (Fig. 3b). The SRA domain does not bind the equivalent non-methylated CpG site, but instead it binds the junction between the two DNA duplexes, arranged head-to-head by the crystallographic twofold symmetry, suggesting that the SRA binds unmethylated DNA non-specifically and only the non-heterogeneous end complex is trapped in crystal form. The side chains of Val 451 and His 450 are in close proximity to the first two base pairs, which are unpaired and non-stacked. These observations are reminiscent of some DNA methyltransferases that

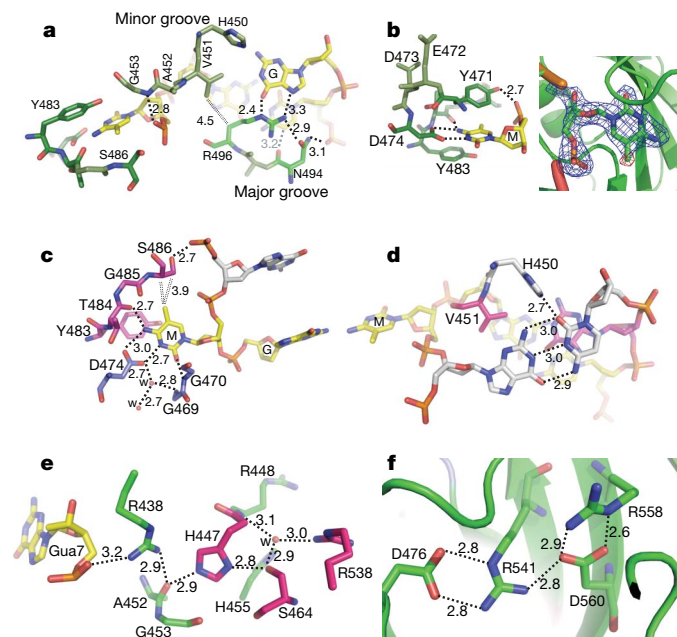


Figure 2 | Details of SRA–DNA interactions. **a**, The 5mC•G base pair is shown in the front, and the adjoining G•C base pair is in the back. **b**, Planar stacking contacts of the extrahelical 5mC with Y471 and Y483 (left image). Omit electron densities, contoured at 4σ and 5σ above the mean, respectively, are shown for omitting 5mC (blue) or the methyl group (red) (right image). **c**, The hydrogen bond interactions with the polar atoms of 5mC. The double-dotted lines indicate van der Waals contacts with the methyl group of ring carbon C5. **d**, H450 forms a hydrogen bond from the minor groove side with cytosine of G•C pair at position 5 (see Fig. 1a). **e**, Network of internal polar interactions centred on residues H447 and S464. Gua, guanine. **f**, Network of internal charged interactions centred on residues R541 and D560. Distances are shown in angstroms.

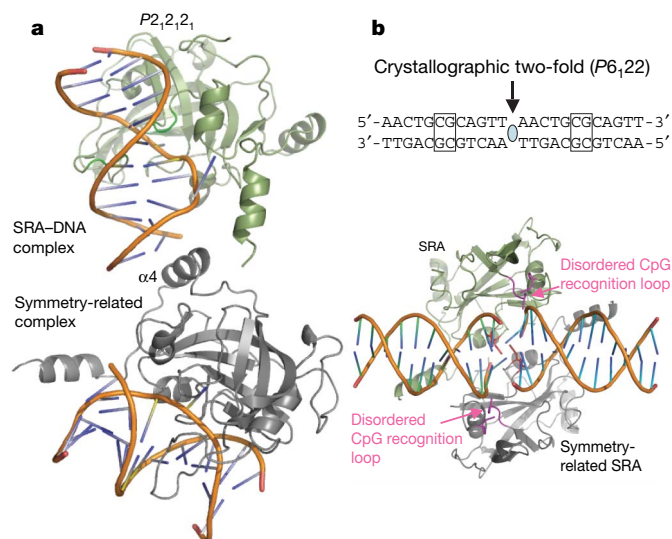


Figure 3 | Structure of the SRA-DNA specific (P2₁2₁2₁) and non-specific (P6₁22) complexes. **a**, The crystal packing interactions in the P2₁2₁2₁ space group may stabilize the DNA duplex. The extra contacts are made by the helix α4 of a symmetry-related molecule. These interactions are not present in the P4₁2₁2 complex. **b**, A 12-base-pair duplex containing a single, centrally located unmethylated CG site was used to generate a non-specific complex with the SRA domain. A V451/H450-associated loop approaches the DNA from the minor groove at the junction of two head-to-head DNA molecules.

also bind at the junction of two duplexes^{12–14}, which mimics damaged DNA or altered recognition sites¹⁵.

A role for UHRF1 in DNA methylation control is widely conserved because mutation of the *Arabidopsis* homologue VIM1 (also known as ORTH2) reduces CG methylation and its SRA domain binds methylated CG sites^{16,17}. In addition, three missense mutations arising in *Arabidopsis* SRA domains are required for binding methylated DNA¹⁸. The corresponding residues in mouse UHRF1 are Ser 464, Glu 472 and Arg 541 (Supplementary Fig. 2). Ser 464 and Arg 541 are involved in intramolecular interactions that probably confer stability to the molecule (Fig. 2e, f). Glu 472 is part of the 5mC-binding pocket, but is not directly involved in interaction with the flipped 5mC (see Fig. 2b).

Base flipping is a conserved mechanism that is widely used by nucleotide-modifying enzymes, including DNA methyltransferases^{19,20}, DNA repair enzymes^{21–23} and RNA modification enzymes²⁴. This mechanism, first discovered in the bacterial 5mC methyltransferase M.HhaI (ref. 19), involves enzyme binding to the DNA and eversion of the target nucleotide so that it projects out of the double helix and into the active-site pocket. The SRA domain is the first non-enzyme sequence-specific DNA-binding protein domain that uses the base flipping mechanism in its interaction with DNA.

There is no apparent sequence or structural similarity between the SRA and the DNA methyltransferase domain (or of DNA repair enzymes). However, the phosphodiester backbone pinching²⁵ caused by extensive protein–phosphate contacts surrounding the flipped nucleotide, the use of two loops to approach DNA from the major and minor grooves simultaneously, and the binding of the flipped base in a concave pocket are analogous (Supplementary Fig. 6)²⁶. Furthermore, enzymes use base flipping to gain access to a DNA base to perform chemistry on it, but the SRA domain probably uses base flipping to increase its protein–DNA interface and to prevent the SRA domain from sliding off the DNA. This may be particularly important for the SRA domain, as its recognition sequence is only two base pairs. The surface area buried at the SRA–DNA interface is approximately 2,500 Å², an ~70% increase from what is buried at the MBD1–DNA interface (1,480 Å²), which does not involve base flipping^{27,28} (Supplementary Fig. 7). The 5mC base flipping by the SRA domain might also provide a more general mechanism to distinguish

the methylated parental strand from the unmethylated daughter strand, an ability particularly important for mismatch repair if an error occurs during DNA replication. Supporting this hypothesis, the expression of ICBP90 (the human orthologue of UHRF1) is deregulated in cancer cells²⁹, and mouse UHRF1-null cells are more sensitive to DNA-damaging agents and DNA replication arrest³⁰. We therefore suggest that the SRA–DNA interaction (through recognition and flipping of the 5mC) serves as a placeholder to keep UHRF1 at hemimethylated CpG site where it recruits DNMT1 for maintenance methylation, and perhaps other proteins such as DNA repair enzymes for mismatch repair.

METHODS SUMMARY

We generated a hexahistidine–SUMO (small ubiquitin-like modifier)-tagged construct containing mouse UHRF1 residues 419–628 (pXC666). The fusion protein was cleaved and the SRA domain was crystallized with DNA containing either a hemimethylated CpG site or a non-specific DNA. The structures were solved by molecular replacement.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 March; accepted 23 July 2008.

Published online 3 September 2008.

- Bostick, M. *et al.* UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
- Sharif, J. *et al.* The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
- Cheng, X. & Blumenthal, R. M. Mammalian DNA methyltransferases: a structural perspective. *Structure* **16**, 341–350 (2008).
- Goll, M. G. & Bestor, T. H. Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* **74**, 481–514 (2005).
- Chen, T. & Li, E. Establishment and maintenance of DNA methylation patterns in mammals. *Curr. Top. Microbiol. Immunol.* **301**, 179–201 (2006).
- Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M. C. & Leonhardt, H. Recruitment of DNA methyltransferase I to DNA repair sites. *Proc. Natl Acad. Sci. USA* **102**, 8905–8909 (2005).
- Jeltsch, A. Molecular enzymology of mammalian DNA methyltransferases. *Curr. Top. Microbiol. Immunol.* **301**, 203–225 (2006).
- Unoki, M., Bronner, C. & Mousli, M. A concern regarding the current confusion with the human homolog of mouse Np95, ICBP90/UHRF1. *Radiat. Res.* **169**, 240–244 (2008).
- Achour, M. *et al.* The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression. *Oncogene* **27**, 2187–2197 (2008).
- Ooi, S. K. & Bestor, T. H. Cytosine methylation: remaining faithful. *Curr. Biol.* **18**, R174–R176 (2008).
- Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y. & Shirakawa, M. Recognition of hemimethylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* doi:10.1038/nature07249 (this issue).
- Horton, J. R., Liebert, K., Bekes, M., Jeltsch, A. & Cheng, X. Structure and substrate recognition of the *Escherichia coli* DNA adenine methyltransferase. *J. Mol. Biol.* **358**, 559–570 (2006).
- Horton, J. R., Liebert, K., Hattman, S., Jeltsch, A. & Cheng, X. Transition from nonspecific to specific DNA interactions along the substrate-recognition pathway of dam methyltransferase. *Cell* **121**, 349–361 (2005).
- Yang, Z. *et al.* Structure of the bacteriophage T4 DNA adenine methyltransferase. *Nature Struct. Biol.* **10**, 849–855 (2003).
- Klimasauskas, S. & Roberts, R. J. M. HhaI binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res.* **23**, 1388–1395 (1995).
- Woo, H. R., Pontes, O., Pikaard, C. S. & Richards, E. J. VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev.* **21**, 267–277 (2007).
- Johnson, L. M. *et al.* The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr. Biol.* **17**, 379–384 (2007).
- Malagnac, F., Bartee, L. & Bender, J. An *Arabidopsis* SET domain protein required for maintenance but not establishment of DNA methylation. *EMBO J.* **21**, 6842–6852 (2002).
- Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76**, 357–369 (1994).
- Cheng, X. & Roberts, R. J. AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Res.* **29**, 3784–3795 (2001).
- Yang, C. G. *et al.* Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* **452**, 961–965 (2008).
- Min, J. H. & Pavletich, N. P. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* **449**, 570–575 (2007).
- Parker, J. B. *et al.* Enzymatic capture of an extrahelical thymine in the search for uracil in DNA. *Nature* **449**, 433–437 (2007).

24. Lee, T. T., Agarwalla, S. & Stroud, R. M. A unique RNA fold in the RuvA–RNA–cofactor ternary complex contributes to substrate selectivity and enzymatic function. *Cell* **120**, 599–611 (2005).
25. Werner, R. M. *et al.* Stressing-out DNA? The contribution of serine–phosphodiester interactions in catalysis by uracil DNA glycosylase. *Biochemistry* **39**, 12585–12594 (2000).
26. Cheng, X. & Blumenthal, R. M. Finding a basis for flipping bases. *Structure* **4**, 639–645 (1996).
27. Ohki, I. *et al.* Solution structure of the methyl–CpG binding domain of human MBD1 in complex with methylated DNA. *Cell* **105**, 487–497 (2001).
28. Ho, K. L. *et al.* MeCP2 binding to DNA depends upon hydration at methyl–CpG. *Mol. Cell* **29**, 525–531 (2008).
29. Mousli, M. *et al.* ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. *Br. J. Cancer* **89**, 120–127 (2003).
30. Muto, M. *et al.* Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. *J. Biol. Chem.* **277**, 34549–34555 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. M. Blumenthal for critical comments. The Emory University School of Medicine supported the use of SER-CAT beamlines. This work was supported by grant GM049245 to X.C. from the National Institutes of Health (NIH). Work in the Jacobsen laboratory is funded by the NIH grant GM060398. M.B. is funded by NIH-NSRA Fellowship number CA1263022. S.E.J. is an Investigator of the Howard Hughes Medical Institute and X.C. is a Georgia Research Alliance Eminent Scholar.

Author Information The X-ray structures (coordinates and structure factor files) of mouse UHRF1 SRA with bound DNA have been submitted to PDB under accession numbers 2ZO0 (P2₁2₁2₁), 2ZO1 (P4₁2₁2) and 2ZO2 (P6₁22), respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to X.C. (xcheng@emory.edu).

METHODS

Expression and purification. We generated a hexahistidine–SUMO-tagged construct³¹ containing mouse UHRF1 residues 419–628 (pXC666). Protein was expressed in *Escherichia coli* BL21(DE3)-Gold cells (Stratagene) with the RIL-Codon plus plasmid. Expression cultures were grown at 25 °C in autoinduction medium³². Cells were lysed as a 20% (v/v) suspension in 50 mM sodium phosphate, pH 7.4, 300 mM NaCl, 5% (v/v) glycerol, 0.1% 2-mercaptoethanol and 40 mg ml⁻¹ phenylmethylsulphonyl fluoride by two passes through an ice-cold French pressure cell press. The lysate was clarified by centrifugation twice at 50,000g for 30 min. Hexahistidine–SUMO fusion proteins were isolated on a nickel-charged chelating column (GE Healthcare). After imidazole elution, the fusion protein was cleaved with Ulp1 protease³³ at 25 U ml⁻¹ with 16 h dialysis at room temperature (approximately 20 °C). Only two extraneous N-terminal amino acids (HisMet) were left as a result of the NdeI restriction site. The protein was further purified by ion exchange (HiTrap-S) and gel filtration chromatography (Superdex-75, GE Healthcare). A maximum yield of 2.8 mg l⁻¹ for the UHRF1 SRA domain was obtained. The domain is monomeric and showed little aggregation during gel filtration in the presence of 5% (v/v) glycerol and 0.1% 2-mercaptoethanol.

Crystallography. For co-crystallization with DNA, the SRA was concentrated to approximately 42 mg ml⁻¹ in 20 mM HEPES, pH 7.0, 150 mM NaCl, 0.1% 2-mercaptoethanol and 5% glycerol, and incubated with an annealed oligonucleotide duplex (5'-TCCATGCGCTGAC-3' and 5'-GTCAGMGCATGG-3', where M = 5mC) at a 1:1.2 ratio for 30 min before setting up crystallization drops. The crystals appeared in various morphologies, sometimes in the same drops. Diamond-shaped (P₄2₁2) and needle-shaped (P₂1₂1₂) crystals were grown under a 20% (v/v) polyethylene glycol 3350, 0.4 M NaCl condition. X-ray diffraction data were collected, at the SER-CAT beamline, and processed with HKL2000³⁴, from crystals cryoprotected by the mother liquor supplemented with 40% ethylene glycerol and allowing crystals to soak for several minutes.

For crystallization with unmethylated oligonucleotide, SRA protein (~32.5 mg ml⁻¹) and annealed DNA (5'-AACTGCGCAGTT-3') were mixed at a 1:1.5 ratio and were incubated for 30 min before setting up crystallization drops. Diamond-shaped crystals grown under 20% (v/v) polyethylene glycol 3350 or 10000, 0.2–0.4 M NaCl and 0.1 M MES (pH 5.8–6.2 or without buffer). However, only one crystal (out of 40–50 crystals screened) diffracted X-ray strong enough and allowed us to collect a complete data set to 3.09 Å resolution.

During the course of study, the Structural Genomics Consortium deposited coordinates of the apo structure of the human SRA domain (PDB accession number 3BI7), which encouraged us to focus on the SRA–DNA complex. We

used the molecular replacement program PHASER³⁵ to obtain crystallographic phases. The resulting electron density map for DNA and the structured portion of the CpG recognition loop (residues 483–496) was easily interpretable, using the model-building program O³⁶. CNS³⁷ scripts were used for refinement, and the statistics shown in Supplementary Table 1 were calculated for the entire resolution range. The *R*_{free} and *R*_{work} values were calculated for 5% (randomly selected) and 95%, respectively, of observed reflections. The structures of P₂1₂1 and P₄2₁2 were solved, built and refined independently. For the non-specific complex structure, discontinued densities do exist but we were not able to unambiguously distinguish between solvent molecules and disordered protein densities. The assignment of solvent molecules to these residual densities would reduce values of *R*_{factor} and *R*_{free}; however, we took a conservative approach without including such solvent molecules in the final model.

31. Lan, F. *et al.* Recognition of unmethylated histone H3 lysine 4 links BHC80 to LSD1-mediated gene repression. *Nature* **448**, 718–722 (2007).
32. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
33. Malakhov, M. P. *et al.* SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. *J. Struct. Funct. Genomics* **5**, 75–86 (2004).
34. Otwinowski, Z., Borek, D., Majewski, W. & Minor, W. Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A* **59**, 228–234 (2003).
35. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60**, 432–438 (2004).
36. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
37. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).

CORRIGENDUM

doi:10.1038/nature07346

Functional metagenomic profiling of nine biomes

Elizabeth A. Dinsdale, Robert A. Edwards, Dana Hall, Florent Angly, Mya Breitbart, Jennifer M. Brulc, Mike Furlan, Christelle Desnues, Matthew Haynes, Linlin Li, Lauren McDaniel, Mary Ann Moran, Karen E. Nelson, Christina Nilsson, Robert Olson, John Paul, Beltran Rodriguez Brito, Yijun Ruan, Brandon K. Swan, Rick Stevens, David L. Valentine, Rebecca Vega Thurber, Linda Wegley, Bryan A. White & Forest Rohwer

Nature 452, 629–632 (2008)

In this Letter, functional diversity and evenness were calculated using log₁₀, but it has been drawn to our attention that the calculations should have been made with natural logs (Table 2). Recalculation does not change the relative levels of diversity; however, this reanalysis showed that a mistake was made in the original evenness calculation. Corrected values are shown in Table 1 below.

Table 1 | Functional metagenomic diversity calculated using the natural log (*H'* max 6.483) and corrected evenness (± s.e.m.)

Biome	Diversity natural log		Evenness	
	Microbial	Viral	Microbial	Viral
Subterranean	5.511 (± 0.098)		0.881 (± 0.008)	
Hypersaline	5.435 (± 0.043)	4.699 (± 0.169)	0.892 (± 0.006)	0.893 (± 0.013)
Marine	5.326 (± 0.134)	4.979 (± 0.188)	0.875 (± 0.021)	0.856 (± 0.015)
Freshwater	5.595 (± 0.015)	4.790 (± 0.155)	0.899 (± 0.004)	0.888 (± 0.015)
Coral	3.991 (± 0.363)	5.271 (± 0.130)	0.748 (± 0.040)	0.901 (± 0.015)
Microbialites	5.544 (± 0.059)	4.014 (± 0.460)	0.903 (± 0.011)	0.830 (± 0.085)
Fish	5.634 (± 0.007)	5.615 (± 0.001)	0.908 (± 0.002)	0.913 (± 0.003)
Terrestrial animals	5.590 (± 0.040)	4.642 (± 0.400)	0.886 (± 0.006)	0.965 (± 0.009)
Mosquito		5.514 (± 0.060)		0.875 (± 0.011)

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

It takes time, effort and no small amount of talent to successfully complete postgraduate and postdoctoral studies. But the next step in a scientific career — a faculty position — comes with a fresh round of challenges. Unlike mastering a new lab technique, there is no ready-made instruction manual for learning how to manage and run a lab for the first time.

This issue was addressed late last month by Matthias Haury, training manager at the European Molecular Biology Laboratory in Heidelberg, Germany, when he spoke at a *Naturejobs* careers conference in London. Haury offered advice on how to approach the setting up of a new lab. First and foremost, he said, you need to think about the big picture. Examine the academic and science enterprise structure of the country where you plan to base your lab — funding and policy issues may affect a lab's long-term future and near-term struggles. In addition, Haury said, you need to vet all of the elements that could affect the lab's productivity, from grant and equipment availability to core facilities — even water quality is worth checking as it could affect experimental outcomes.

Another key step, Haury noted, is to draw up your budget before you begin. This needs to include consumables and small lab equipment, as well as having some room for contingencies. Then you need to work out who to recruit. Quality rather than quantity of staff is important — although Haury said that he has had more success with applicants with creative CVs and unconventional backgrounds rather than just stellar grades. Above all, personal rapport is key.

Lab-management courses can be helpful, such as those offered by the European Molecular Biology Organization or the Howard Hughes Medical Institute (see *Nature* **451**, 740–741; 2008). But Haury emphasized the route to success involves maximizing your “personal impact factor” by speaking at meetings, publishing and writing reviews. The better this impact factor, the better your networking opportunities and grant possibilities — and the better the chances that you will attract superb staff.

Gene Russo is editor of *Naturejobs*.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com

Natureevents:
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com

UK Corporate:
Nils Moeller (4953)

Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)

France/Switzerland/Belgium:

Muriel Lestringuez (4994)

Scandinavia/Spain/Portugal/Italy:

Evelina Rubio-Hakansson (4973)

North Germany/The Netherlands/Eastern

Europe: Reya Silao (4970)

South Germany/Austria:

Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:

Stephen Russell

To send materials use London address above.

Tel: +44 (0) 20 7843 4816

Fax: +44 (0) 20 7843 4996

e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock

Naturejobs online production: Dennis Chu

US Head Office, New York

75 Varick Street, 9th Floor,

New York, NY 10013-1917

Tel: +1 800 989 7718

Fax: +1 800 989 7103

e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India

Vikas Chawla (+91 1242881057)

e-mail: v.chawla@nature.com

Japan Head Office, Tokyo

Chiyoda Building, 2-37 Ichigayatamachi,

Shinjuku-ku, Tokyo 162-0843

Tel: +81 3 3267 8751

Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:

Ayako Watanabe (+81 3 3267 8765)

e-mail: a.watanabe@natureasia.com

Business Development Manager, Greater

China/Singapore:

Gloria To (+852 2811 7191)

e-mail: g.to@natureasia.com

MOVERS

James Pendlebury, chief executive, Forest Research, Alice Holt Lodge, Farnham, UK.



2007-08: Head of specialist advisers, Forestry Commission, Edinburgh, Scotland

2004-06: Business policy adviser, Forestry Commission Scotland, Edinburgh, Scotland

2002-04: Partnership manager, Forest Enterprise Scotland, Inverness, Scotland

James Pendlebury is not a typical materials scientist. He studies the properties of wood rather than metal alloys: a niche that has helped sprout a gratifying career path.

Trained in forestry, Pendlebury has travelled the world, largely helping industry assess local species' unique properties. That has given him a global view of forestry concerns that he says will help him direct the UK Forestry Commission's research agenda, focused on climate change, bioenergy and sustainable forest management.

His roots, however, are in research. "The essence of wood science is understanding the material itself — the structural and chemical properties — that make each species unique," he says. With industry support, he pursued a PhD at the University of Aberdeen in Scotland, studying how Malaysian tropical hardwoods could be processed for use in construction. His PhD adviser, Alan Petty, says such an industry-funded project was unusual at the time. But it gave Pendlebury an early appreciation for applying research to industry.

Pendlebury developed an expertise for finding new uses for timber species as he travelled the world. As a postdoc at the University of Maine in Orono he investigated whether local spruce species could replace imported pine as telegraph poles. He investigated wood preservation and pulp properties for the South African government. Then he worked on preservation techniques at the TNO Centre for Timber Research in the Netherlands and in New Zealand.

He returned to Britain, after eight years, as a timber consultant for TRADA Technology Limited, a research consultancy near London. He then eagerly took charge of a small non-profit forest management company called Highland Birchwoods in Ross-shire, Scotland. "There I learned how important it is to link local economic benefit with environmental gain," he says. Six years ago, he joined the Forestry Commission, helping to start large-scale forestry projects in Scotland. In his new position, he will oversee the nascent Centre for Forestry and Climate Change, opening next year. Aims include identifying potential pest risks and alternative forestry practices.

"Adaptation to and mitigation of climate change will be key," says Pendlebury, adding that forestry's role is not only about restoring forest cover, but promoting wood energy from renewable timber sources. Exploring forests' carbon sequestering potential, says Petty, is likely to be a major objective for Pendlebury.

Virginia Gewin

NETWORKS & SUPPORT

Developing generic skills

At the University of Dundee in Scotland, the generic-skills programme has a rather specific aim: to prepare young scientists for a challenging employment landscape. Building on its success, the programme has added a series of lunchtime workshops with truncated versions of some courses that emphasize one-to-one instruction in subjects ranging from tweaking a CV to matching skills and experience to jobs outside academia.

More than 50% of Dundee's life-sciences graduate students and 40% of its life-science postdocs have taken the full-blown courses in the two years they have been available, says training coordinator Christine Milburn.

The courses emerged following the 2002 Roberts report, which recommended ways to infuse Britain's science, engineering and technology pipeline with talent. It emphasized training in soft-skills areas: research environment, research management, communication skills, personal effectiveness, networking, team working and career management. The Dundee programme includes grant writing, presenting papers at conferences and managing projects.

The most heavily subscribed course emphasizes communicating science to non-specialists; its popularity

reflects the growth of interdisciplinary research, where chemists, physicists and biologists need to find a common language. Communications coursework is also working to reach the non-scientific community.

The programme works with the university's postdoc and graduate-student organizations to pick the curriculum and refine content, Milburn says. Feedback has been especially positive for courses that encourage graduate students and postdocs to 'map' their skills and interests to different career paths in both academia and industry, then seek training to fill any gaps.

As a result, participants say they don't feel locked into pursuing academic careers; the programme warns them that only a small percentage will land university jobs. It illuminates other pathways, such as media, marketing and technology transfer, says Milburn.

"A lot of people who are taking part in this are looking at a new direction," she says. "They feel the university is supporting their career, not just their research." Other Scottish universities are now emulating the programme, including transferable skills courses at the University of St Andrews and the University of Edinburgh.

Paul Smaglik

POSTDOC JOURNAL

Research growing pains

My postdoctoral experience is about to come to an end. After an intense half-year of interviews and negotiations, I have accepted a position as an assistant professor back home in the United States. I'm glad that all the travelling, seminar preparations and in-depth scientific conversations have paid off. But I loathe transitions: research growing pains can be painful. Settling into my postdoc three years ago took time and effort, not only because of the culture shock of being in Israel, but because my new direction, studying the effects of genes on an entire plant, required a rewiring in the way I approach science.

Now I must rewire again. I will be managing a lab, doing less hands-on work. Some say that the postdoc is the best time of a scientist's career. I can see why: there is freedom of research without the stresses of grant writing; mentoring is more fun and less business; there's less politics and bureaucracy.

Is it all downhill from here? How high will the stress level climb now that I am responsible for grant writing, teaching, thesis committees, extensive mentoring and striving for the long-term goal of tenure? I'm trying not to obsess too much about these issues, because I fear it may prevent me from making that first step onto the plane back home. But the reality is that I am excited in so many ways, and this new energy has already started to dull the pain of the transition.

Zachary Lippman is a postdoctoral fellow at the Hebrew University of Jerusalem's faculty of agriculture.

The picture of Oscar X

An image of the future.

**Anna Batistatou and
Konstantinos Charalabopoulos**

Here I am, an old man, writing the weirdest and probably the last death certificate of my medical career. I have been practising internal medicine for more than 50 years and I am now a well-respected physician, close to retirement and daring to narrate the story of my last patient.

I was in the middle of my career when I first met Oscar X. He was a 55-year-old man of medium height and weight, bald with a small beard. He had complained of pain in his right hip for the previous six months. Initially the pain was mild, but it became stronger, and swelling appeared in the area. X-rays revealed a radiolucent mass with calcifications at the right ilium. MRI did not show extension to the adjacent soft tissues. A biopsy revealed the presence of chondrosarcoma grade 1 with areas of dedifferentiated chondrosarcoma.

The prognosis was dismal.

Oscar was a cultured, highly educated man; a painter. He wanted to know more about the neoplasm that was threatening his ability and life. The monochrome images of the X-rays and MRI scans did not satisfy him. The 'thing' was unimpressive; its deadly potential, muted. After visiting the oncologist, Oscar decided to meet Dr Wilde, the pathologist who had diagnosed his neoplasm. He had a favour to ask: he wanted to see the *colour* of his cancer.

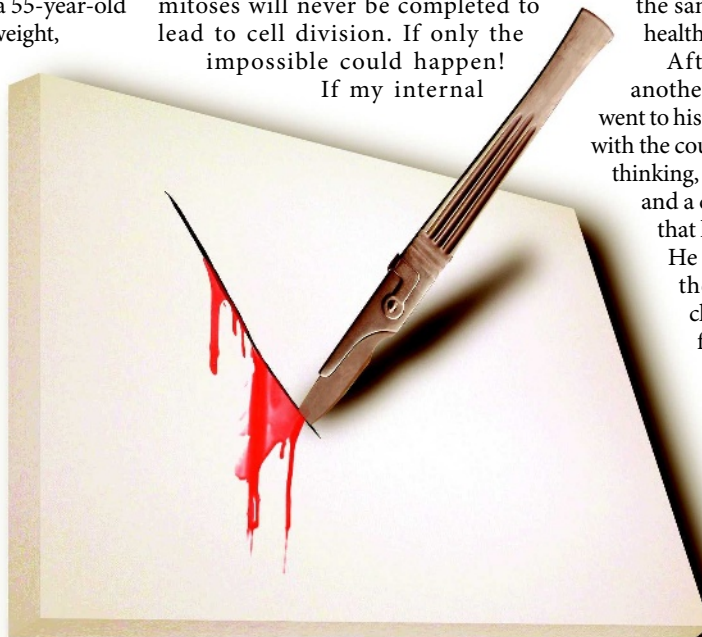
Dr Wilde put the glass slide under her microscope and let him see. She was surprised by his reaction. She considered the neoplasm ugly, with all that nuclear pleomorphism and the abundant atypical mitoses, but to him it was beautiful. He explained that this was the kind of appearance he would expect from his neoplasm. To his eyes, the eyes of a painter, it looked like a piece of art, with spectacular shapes and colours. He asked for a photo.

Next day in his studio he painted a portrait of his neoplasm. He took care to depict every detail: the blue-grey colour of the matrix, the pink fibrous tissue, the

blue pleomorphic nuclei and the atypical mitoses. When he finished he stepped back and blushed with happiness. The painting was so beautiful, so colourful — and so cheerful! But still he could not forget that this neoplasm was going to kill him.

"How sad this is," he thought, "my neoplasm will grow, it will metastasize and finally it will kill me. But this picture will always remain the same. It will not grow further than it is today. These mitoses will never be completed to lead to cell division. If only the impossible could happen!

If my internal



neoplasm would remain stable and the one in the picture would grow and expand. I would give anything for this to happen."

Oscar was a strong-willed man. For reasons he could not explain, he denied all treatment, and agreed only to regular follow-ups with imaging. As for his painting, it became a major hit of his latest exhibition, but he refused to sell it, even when offered high prices. Instead, he hung the painting in his bedroom.

One day, coming back late from a night out, Oscar casually glanced at the portrait. He was startled, as if he saw it for the first time. He hesitated and looked again. Under the dim light he noticed that the atypical mitoses were not where he had drawn them so carefully. Instead, others had appeared, elsewhere in the painting. New blood vessels, with plump endothelial and red blood cells, were also apparent. How could this happen? Suddenly he remembered the wish he'd made when he

finished the painting. Was it possible that it had come true? Being a man of reason he knew that such things did not happen. He hid the painting behind heavy curtains and did not look at it again.

After some time he came to the hospital for his follow-up. And, to our surprise, his disease had remained stable. X-rays and an MRI scan showed that the neoplasm had remained the same size, and there were no metastases. Each follow-up produced the same results, and Oscar looked healthy.

After returning home from another visit to the hospital, Oscar went to his room to undress. Being happy with the course of his disease and without thinking, he looked at the heavy curtains and a chill came over him. He knew that he had to look at the painting. He took a deep breath and drew the curtains. The picture had changed and the view was terrifying. Now it was three-dimensional, protruding from the surface of the canvas. It was a real mass. Also new nests of similar cells were obvious at the wall adjacent to the painting. There it was, the very death represented by neoplastic cells. But the irony was that the canvas that carried it could not die.

From then on Oscar checked the painting regularly. He saw areas of necrosis appear; he saw the metastases grow and give new metastases, like ivy growing on the wall. And he kept growing old and looking healthy.

Several years later, Oscar could not stand the view of death any more. He took a knife and stabbed the painting repeatedly. At the same time he fell to the floor bleeding profusely. He was admitted to our hospital, and he asked for me. Scans revealed a huge mass on the ilium with wide extension to adjacent soft tissues and innumerable metastases to the lungs and liver. How strange! I had seen him only two weeks before and none of this was apparent.

He died a few days later.

Anna Batistatou and Konstantinos Charalabopoulos are at the University of Ioannina, Medical School, Ioannina, Greece, and often contemplate the similarities between science and art.

JACEY